J. Math. Kyoto Univ.
8-2 (1968) 233-278

Branching Markov processes I

By

Nobuyuki IKEDA, Masao NAGASAWA and Shinzo WATANABE(*)

(Received May 10, 1968)

Introduction. Let S be a compact Hausdorff space with a countable open base, S^n the n-fold symmetric product of S, $\boldsymbol{S}=\bigcup_{n=0}^{\infty} S^n$ the topological sum of S^n, where $S^0=\{\partial\}$, ∂ an extra point, and $\hat{\boldsymbol{S}}=\boldsymbol{S}\cup\{\Delta\}$ the one-point compactification of $\boldsymbol{S}$. The purpose of this paper is to investigate a class of semi-groups $\{\boldsymbol{T}_t;\ t\geq 0\}$ of linear operators defined on the space $\boldsymbol{B}(\hat{\boldsymbol{S}})$ of bounded measurable functions on $\hat{\boldsymbol{S}}$ with a special property, which will be called the *branching property*;

$$\boldsymbol{T}_t\hat{f}(\boldsymbol{x})=\widehat{(\boldsymbol{T}_t\hat{f})|_S}(\boldsymbol{x}),\quad \boldsymbol{x}\in\boldsymbol{S},\ \hat{f}\in\boldsymbol{B}(S), \tag{1}$$

where $\wedge$ is a mapping from $\boldsymbol{B}(S)$, the space of bounded measurable functions on S, to $\boldsymbol{B}(\hat{\boldsymbol{S}})$ defined by

$$\hat{f}(\boldsymbol{x})=\begin{cases}\prod_{j=1}^{n} f(x_j), & \text{when } \boldsymbol{x}=[x_1, x_2, \cdots, x_n],\\ 1 & \text{, when } \boldsymbol{x}=\partial,\\ 0 & \text{, when } \boldsymbol{x}=\Delta.\end{cases} \tag{2}$$

When the semi-group $\boldsymbol{T}_t$ is positivity preserving and contraction, there corresponds a Markov process on $\hat{\boldsymbol{S}}$ with the semi-group by the general theory of Markov processes. We shall call the Markov process a *branching Markov process*. Branching processes are investigated by many authors as a mathematical model for the population

---

* The first and the third authors were supported in part by contract USPHFNIH 10452-04 through Stanford University and the second author, by NSF Grant GP 4867 through Cornell University.

growth of particles (cf., e.g., Harris [8]).

This paper will consist of three parts. In the first part we shall give some equivalent formulations of the branching property. The most important one is the equivalence between the branching property and the property B. III (Theorem 1.2). Roughly speaking, the property B. III is equivalent to say that if there are n-particles they move independently of each other before the splitting (branching) time, and when they split into m-particles this happens only through the event that just one of the original n-particles splits into $m-(n-1)$-particles and the other $(n-1)$-particles continue the same motion as before. Several versions of the property which is equivalent to the Property B. III are adopted as a definition of branching processes by some authors, but the equivalence between the branching property (1) and the Property B. III, as far as the authors know, has not been proved in full generality. This equivalence will play an important role especially in constructing branching Markov processes from given fundamental quantities. The equivalence will be proved in part I, while the construction itself will be treated in part II. Using the strong Markov property of the branching Markov process, one can easily see that there are two fundamental quantities which will uniquely determine the process. They are the *non-branching part* X^0 (Definition 1.2) of the branching Markov process and the *branching law* $\pi(x, dy)$ (Definition 1.3). The first one is a Markov process on $S\cup\{\Delta\}$ with Δ as the terminal point (it must be remarked that the state space is not $\boldsymbol{S}$ but S), which describes the behavior of a particle before its first splitting. The second one is a stochastic kernel on $S\times\widehat{\boldsymbol{S}}$ such that $\pi(x, S)=0$ for every $x\in S$, which governs the law of the splitting. We shall prove that a large class of branching processes is uniquely determined by X^0 and π (such a branching Markov process will be called an (X^0, π)-branching Markov process).

In part II, we shall first give a general theorem of piecing out for Markov processes. This theorem will provide a simple way to piece out the sample functions of a given Markov process by an instantaneous distribution (Definition 2.1). Thus, the problem of construction for branching Markov processes is reduced to construct a Markov process on $\hat{S}$ from the given non-branching part X^0 (a Markov process on S) and an instantaneous distribution from the given branching law (Theorem 3.5), and prove that the constructed Markov process by the piecing out theorem has the property B. III. This is, however, an immediate consequence of the way of construction of the non-branching part on $\hat{S}$ and the instantaneous distribution. We shall give several examples of branching Markov processes in part II. To do this, it is sufficient to specify non-branching parts and branching laws. We shall give there related fundamental equations for these processes, while the general form will be derived in part I for (X^0, π)-branching Markov processes. These are: a linear integral equation on the large state space $\hat{S}$ of renewal type

$$(3) \qquad u(t, \boldsymbol{x}) = \boldsymbol{T}_t^0 f(\boldsymbol{x}) + \int_0^t \int_{\hat{S}} \psi(\boldsymbol{x};\, ds\, dy) u(t-s, \boldsymbol{y}), \qquad \boldsymbol{x} \in \hat{S},\ t \in [0, \infty),$$

where $f \in \boldsymbol{B}(\hat{S})$; and a non-linear integral equation on S

$$(4) \qquad u(t, x) = T_t^0 f(x) + \int_0^t \int_S K(x;\, ds\, dy) F[y;\, u(t-s)], \qquad x \in S,\ t \in [0, \infty),$$

where $f \in \boldsymbol{B}(S)$ and

$$F[x;\, g] = \int_{\hat{S}} \pi(x, dy)\, \hat{g}(y), \quad g \in \boldsymbol{B}(S).$$

We shall call (3) and (4) M-equation and S-equation respectively. These equations are defined in terms of the non-branching part X^0 and the branching law π only.

A detailed discussion for the equations will be given in part III, which may be understood as an analytical version of part II. We shall construct a (minimal) solution of the M-equation, following Moyal [33], and show that it defines a semi-group with the branching property (1). A solution $u(t, x)$ of the S-equation can be constructed by the usual method of successive approximation and from the solution one can define a semi-group $\boldsymbol{T}_t$ on $\boldsymbol{B}(\widehat{S})$ putting $\boldsymbol{T}_t\widehat{f}(\boldsymbol{x})=\prod_{j=1}^{n}u(t, x_j)$, when $\boldsymbol{x}=[x_1, x_2, \cdots, x_n]$. Clearly it has the branching property. Thus these constructions provide two independent ways of analytically constructing an (X^0, π)-branching process. On the other hand, it is easy to show that there is an intimate relation between the solutions of the S-equation and M-equation: that is, the minimal solution of the M-equation provides the one for the S-equation, while a solution of the M-equation can be constructed from the solution of the S-equation. In other words, the M-equation is a *linearization* of the S-equation. By this relation we can investigate the solution of the S-equation (a *non-linear* semi-group on $\boldsymbol{B}(S)$) in terms of the M-equation (a *linear* semi-group on $\boldsymbol{B}(\widehat{S})$).

Assuming some regularity conditions, we shall discuss the infinitesimal generator of a branching semi-group. In doing so, we shall derive two fundamental differential equations: the *backward equation* and the *forward equation.* The backward equation is a quasi-linear evolution equation. Such a class of equations was considered in, e.g., Kolmogoroff-Dmitriev [24] and Bartlett [1] for some of the simplest cases, and by Itô-McKean [19] and Skorohod [43] for branching processes with diffusing particles. In the case of branching Brownian motion, this class of non-linear differential equations is a particular case of the equations discussed by Kolmogoroff, Petrovsky and Piscounoff [25]. The forward equation is a linear evolution equation involving functional derivatives. We shall prove the uniqueness of solutions of the forward equation and apply it to give another proof of the branching property of the minimal solution of the M-equation. Finally, in Chapter IV, we shall discuss the equation related to the number of particles. In particular, we shall see that the first moment defines a nonnegative but not necessarily contraction semi-group. A probabilistic treatment of such a semi-group was treated by Hunt [10] and Knight [23]. A branching Markov process also seems to provide one of the natural and "nice" model of creation of mass.

In Chapter V, we shall discuss transformations of branching Markov processes; i.e., operations on a branching Markov process to get a new branching Markov process. Some interesting examples will be given for killing, drift and harmonic transformation.

The many results of this work have been previously published without proofs in Ikeda, Nagasawa and Watanabe [12], [13], [14], [15], [16], and [17].

We would like to express our deep gratitude to Thomas H. Savits for his kind help in preparing the manuscript.

0. Preliminaries

Necessary facts on Markov processes will be summarized in §0.1. In §0.2, we shall discuss the symmetric product spaces and their direct sum of a given compact metrizable space and define several operations on functions which will play an important role in the future discussions.

§0.1. Markov processes

Let E be a locally compact Hausdorff space with a countable open base, and let $\mathcal{B}(E)$ be the topological Borel field of E, i.e. the smallest Borel field containing all open sets of E. A Markov process X on E is a collection

$$X=(\Omega, \mathcal{B}_t, P_x, X_t(\omega), \theta_t)$$

of the following objects which satisfy the axioms (M. 1)–(M. 4) given below.

(1) Ω is an abstract space,

(2) $\mathcal{B}_t$, $t\in[0, \infty)$ is a family of increasing Borel fields on Ω (we denote the smallest Borel field containing all $\mathcal{B}_t$ by $\mathcal{B}_\infty=\bigvee_{t>0}\mathcal{B}_t$),

(3) $X_t(\omega)$, $t\in[0, \infty)$, $\omega\in\Omega$ is a function $(t, \omega)\in[0, \infty)\times\Omega\to X_t(\omega)\in E$ *adapted to* $\mathcal{B}_t$ i.e. for fixed $t\in[0, \infty)$, the mapping $\omega\to X_t(\omega)$ is measurable $(\Omega, \mathcal{B}_t)\to(E, \mathcal{B}(E))$,

(4) θ_t, $t\in[0, \infty)$, is a system of mappings from Ω to Ω,

(5) P_x, $x\in E$, is a system of probability measures on $(\Omega, \mathcal{B}_\infty)$.

Now, we introduce other Borel fields

$$\mathcal{N}_t=\sigma(\Omega, \mathcal{B}(E);\ X_s(\omega),\ s\leq t)^{(1)} \tag{0.1}$$

and

$$\mathcal{N}_\infty=\bigvee_{t>0}\mathcal{N}_t=\sigma(\Omega, \mathcal{B}(E);\ X_s(\omega),\ s\in[0,\infty)). \tag{0.2}$$

Clearly $\mathcal{N}_t\subset\mathcal{B}_t$ and $X_t(\omega)$ is adapted to $\mathcal{N}_t$.

Definition 0.1. A collection $X=(\Omega, \mathcal{B}_t, P_x, X_t, \theta_t)$ is called a *Markov process on E* if it satisfies the following:

(M.1) $X_{t+h}(\omega)=X_t(\theta_h\omega)$, for all $\omega\in\Omega$ and $t, h\in[0,\infty)$,

(M.2) $P_x[\omega;\ X_0(\omega)=x]=1$, for all $x\in E$,

(M.3) For every $t\in[0,\infty)$ and $A\in\mathcal{B}(E)$, $P_x[\omega;\ X_t(\omega)\in A]$ is a $\mathcal{B}(E)$- measurable function of x,

(M.4) $P_x[X_{t+s}(\omega')\in A\,|\,\mathcal{B}_t]=P_{X_t(\omega)}[X_s(\omega')\in A]$ *a.a.* $\omega(P_x)$ for all t, $s\in[0,\infty)$ and $A\in\mathcal{B}(E)$.

The following statements are easy consequences of the definition.

(i) θ_t is measurable $(\Omega, \mathcal{N}_\infty)\to(\Omega, \mathcal{N}_\infty)$.

(ii) (M.3) and (M.4) are equivalent to the following two conditions: For any bounded $\mathcal{N}_\infty$-measurable function $F(\omega)$,

(M.3)′ $$E_x[F(\omega)]\equiv\int_\Omega F(\omega)P_x[d\omega] \text{ is } \mathcal{B}(E)\text{-measurable in } x$$

and

(M.4)′ $$E_x[F(\theta_t\omega')\,|\,\mathcal{B}_t]=E_{X_t(\omega)}[F],\quad a.a.\ \omega(P_x).$$

Now we shall put

$$\overline{\mathcal{B}}(E)=\bigcap_{\mu\in\mathfrak{M}}\overline{\mathcal{B}}^\mu(E), \tag{0.3}$$

where $\mathfrak{M}$ is the set of all probability measures on $(E, \mathcal{B}(E))$ and $\overline{\mathcal{B}}^\mu(E)$ is the completion of $\mathcal{B}(E)$ by $\mu\in\mathfrak{M}$ and put

1) If $\{f_\alpha\}$ is a family of functions from Θ to a measurable space $(S, \mathcal{B})$, then $\sigma(\Theta, \mathcal{B};\ f_\alpha)$ is the smallest Borel field on Θ with respect to which all f_α: $\Theta\to(S, \mathcal{B})$ are measurable.

(0.4) $$\overline{\mathcal{B}}_t=\bigcap_{x\in E}\overline{\mathcal{B}}_t^{P_x},\ 0\leqq t\leqq+\infty.$$

If X is a Markov process, then (M. 3) and (M. 4) are still valid when we replace $\mathcal{B}(E)$ and $\mathcal{B}_t$ by $\overline{\mathcal{B}}(E)$ and $\overline{\mathcal{B}}_t$ respectively.

From now on, unless otherwise stated, all the Markov processes we are considering are supposed to satisfy the following condition (R.C) of right continuity;

(R.C). *For every $\omega\in\Omega$ the mapping*

$$[0,\infty)\ni t\to X_t(\omega)\in E$$

is right continuous.

Then it follows that $X_t(\omega)$ is *progressively measurable*; i.e., for every $u\in[0,\infty)$, $(t,\omega)\to X_t(\omega)$ is measurable $([0,u]\times\Omega,\ \mathcal{B}_{[0,u]}\otimes\mathcal{B}_u)\to(E,\mathcal{B}(E))$[2].

A non-negative random time $T(\omega)$: $\Omega\to[0,\infty]$ is called a *$\mathcal{B}_t$-Markov time* if $\{T\leqq t\}\in\mathcal{B}_t$ for every $t\in[0,\infty)$. For a given $\mathcal{B}_t$-Markov time, we define a Borel field by

(0.5) $$\mathcal{B}_T=\{B\in\mathcal{B}_\infty;\ \text{for every } t\in[0,\infty),\ B\cap\{T\leqq t\}\in\mathcal{B}_t\}$$

(Note that if $T\equiv t$ then $\mathcal{B}_T=\mathcal{B}_t$).

Definition 0.2. A Markov process $X=(\Omega,\mathcal{B}_t,P_x,X_t,\theta_t)$ is called a *strong Markov process* if for every $\mathcal{B}_t$-Markov time T and $A\in\mathcal{B}(E)$,

(S.M) $$P_x[X_{T+t}\in A,\ T<\infty\,|\,\mathcal{B}_T]$$
$$=I_{\{T<\infty\}}(\omega)P_{X_{T(\omega)}(\omega)}[X_t\in A],\ \ a.a.\ \omega(P_x)^{(3)}.$$

Remark 0.1. (i) Since $X_t(\omega)$ is progressively measurable, $X_{T(\omega)}(\omega)$ is $\mathcal{B}_T/\{T<\infty\}$-measurable[4] and $\omega\to\theta_{T(\omega)}\omega$ is measurable $(\{T<\infty\},\ \mathcal{B}_\infty/\{T<\infty\})\to(\Omega,\mathcal{N}_\infty)$, (cf. Meyer [31]).

(ii) The condition (S.M) is equivalent to

---

2) $\mathcal{B}[0,u]$ is the topological Borel field of $[0,u]$. $\mathcal{B}_{[0,u]}\otimes\mathcal{B}_u$ is the usual product Borel field.

3) $I_{\{T<\infty\}}(\omega)$ is the indicator function of the set $\{\omega\colon T(\omega)<\infty\}$.

4) $\mathcal{B}_T/\{T<\infty\}\equiv\{A\in\mathcal{B}_T\colon A\subset\{T<\infty\}\}$: the restriction of $\mathcal{B}_T$ on $\{T<\infty\}$.

(S. M)′ $E_x[I_{\{T<\infty\}}F(\theta_T\omega)\,|\,\mathcal{B}_T]=I_{\{T<\infty\}}E_{X_T}[F]$ *a.a.* $\omega(P_x)$ for every $\mathcal{B}_t$-Markov time T and bounded $\mathfrak{N}_\infty$-measurable function $F(\omega)$. Also the following variant (S.M)″ of (S.M)′ is useful in applications (cf. Dynkin [6]): if $F(\omega, t)$, $(\omega, t)\in\Omega\times[0, \infty)$, is bounded and $\mathfrak{N}_\infty\otimes\mathcal{B}_{[0,\infty)}$-measurable then

(S. M)″ $$E_x[I_{\{T<\infty\}}F(\theta_T\omega,\ T(\omega))\,|\,\mathcal{B}_T]=I_{\{T<\infty\}}\Phi(T,\ X_T)$$

where $\Phi(s, x)=E_x[F(\omega, s)]$. We write $\Phi(T, X_T)$ sometimes as $E_{X_T}[F(\omega, t)]\big|_{t=T}$.

(iii) By Dynkin [6] (p. 102, Theorem 3.12) or T. Watanabe [47], if $(X_t, \mathcal{B}_t)$ is strong Markov, $(X_t, \overline{\mathcal{B}}_t)$ is strong Markov. Hence without loss of generality we may assume $\overline{\mathcal{B}}_t=\mathcal{B}_t$ for every strong Markov process $(X_t, \mathcal{B}_t)$.

(iv) Since $X_t(\omega)$ is progressively measurable, every hitting time T_A for any set $A\in\mathcal{B}(E)$ is a $\overline{\mathcal{B}}_{t+0}$-Markov time where $\overline{\mathcal{B}}_{t+0}=\bigcap_n\overline{\mathcal{B}}_{t+\frac{1}{n}}$ (cf. Meyer [31]). Hence, if $\mathcal{B}_t$ satisfies the condition $\overline{\mathcal{B}}_{t+0}=\mathcal{B}_t$, then every hitting time for a set $A\in\mathcal{B}(E)$ is a $\mathcal{B}_t$-Markov time.

Definition 0.3. A Markov process $x=(X_t, \mathcal{B}_t)$ is called *quasi-left continuous* if for every increasing sequence of Markov times T_n we have

(Q. L. C) $$P_x[\lim_{n\to\infty} X_{T_n}=X_T,\ T<\infty]=P_x[T<\infty]$$

for every $x\in E$, where $T=\lim_{n\to\infty} T_n$.

Definition 0.4. A Markov process $X=(X_t, \mathcal{B}_t)$ is called a *Hunt process* if it satisfies the following conditions:

(H.1) the existence of left limits, i.e., for every $x\in E$, $P_x[X_{t-0}(\omega)\equiv\lim_{s\uparrow t} X_s(\omega)$ exists for all $t>0]=1$,

(H.2) $(X_t, \mathcal{B}_t)$ is strong Markov,

(H.3) $(X_t, \mathcal{B}_t)$ is quasi-left continuous,

(H.4) $\overline{\mathcal{B}}_{t+0}=\mathcal{B}_t$.

We shall define, as usual, the equivalence of two stochastic processes by the coincidence of their finite dimensional joint distributions. Then, by virtue of the Markov property, it is easy to see that two Markov processes X and X' on E are equivalent if and only if $P_x[X_t \in A] = P'_x[X'_t \in A]$ for every $x \in E$ and $A \in \mathscr{B}(E)$. Clearly this is quivalent to saying that X and X' induce the same semi-group on $\boldsymbol{B}(E)$[5], i.e., if $T_t f(x) = E_x[f(X_t)]$ and $T'_t f(x) = E'_x[f(X'_t)]$, $f \in \boldsymbol{B}(E)$, then $T_t = T'_t$.

A point $x \in E$ is called a *trap* if

$$P_y[\{\omega;\ \text{if } X_s(\omega) = x \text{ then } X_t(\omega) = x \text{ for all } t \geqq s\}] = 1 \tag{0.6}$$

for every y. It is sometimes convenient to distinguish a trap $\Delta \in E$ and to consider that when X_t reaches Δ it has terminated its life. Such a point Δ is called the ***terminal point*** of the process X_t and the hitting time $\zeta(\omega)$ for Δ,

$$\zeta(\omega) = \inf\{t;\ X_t(\omega) = \Delta\},^{(6)} \tag{0.7}$$

is called the *life time* of X_t.

Here is one of the most fundamental theorems in the theory of Markov processes (cf. *e.g.* Dynkin [6]). Let $\hat{E} = E \cup \{\Delta\}$ be one point compactification of E[7] and $\hat{\boldsymbol{C}}(E)$ be the set of all continuous functions on $\hat{E}$ such that $f(\Delta) = 0$. $\hat{\boldsymbol{C}}(E)$ is a real Banach lattice with norm $\|f\| = \sup_{x \in \hat{E}} |f(x)|$ and with the usual order. Let $\{T_t, t \geqq 0\}$ be a semi-group of operators on $\hat{\boldsymbol{C}}(E)$ satisfying the conditions

(T. 1) if $1 \geqq f \geqq 0$, then $1 \geqq T_t f \geqq 0$, and

(T. 2) $\|T_t f - f\| \to 0$ where $t \downarrow 0$ for every $f \in \hat{\boldsymbol{C}}(E)$.[8]

Then the theorem reads as follows:

Theorem 0.1. *Let* $\{T_t, t \geqq 0\}$ *be a semi-group satisfying*

---

5) $\boldsymbol{B}(E)$ is the set of all bounded (real valued) $\mathscr{B}(E)$-measurable functions on E.

6) We shall always set $\inf \phi = +\infty$ by convention.

7) When E is already compact, Δ shall be an isolated point of E.

8) Such a semi-group is celled a *strongly continuous, non-negative, contraction semi-group.*

(T. 1) *and* (T. 2). *Then there exists a Hunt process* $X=(\Omega, \mathcal{B}_t, P_x, X_t, \theta_t, \zeta(\omega))$ *on* $\hat{E}=E\cup\{\Delta\}$ *with* Δ *as the terminal point such that*

(0. 8) $$T_tf(x)=E_x[f(X_t)]$$

for all $x\in\hat{E}$, $t\geqq 0$ *and* $f\in\hat{C}(E)$. *Such* X *is unique up to equivalence.*

It is also well known that the above Hunt process can be given in the *canonical form*; that is, (Ω, X_t, θ_t) are given by (i)~(iii) below.

(i) $\Omega=\Omega_{rc}\equiv$the set of all right continuous functions w; $t\in[0, \infty)\to w(t)\in\hat{E}$ such that for some $0\leqq\zeta(w)\leqq+\infty$, $w(t)$ is in E when $t<\zeta(w)$ and $w(t)=\Delta$ for $t\geqq\zeta(w)$.

(ii) $X_t(w)=w(t)$, $w\in\Omega_{rc}$.

(iii) θ_tw is defined by $(\theta_tw)(s)=w(t+s)$.

There are several ways of giving $\mathcal{B}_t$; the following seems to be a standard one. Let $\mathcal{N}_t=\sigma(\Omega_{rc}, \mathcal{B}(\hat{E}); X_s(w), s\leqq t)$, $\mathcal{N}_\infty=\bigvee_{t>0}\mathcal{N}_t$ and

(0. 9) $$\mathcal{F}=\bigcap_{\mu\in\mathfrak{M}}\overline{\mathcal{N}_\infty}^{P_\mu},^{(9)}$$

where P_μ, $\mu\in\mathfrak{M}$, is a probability measure on $(\Omega_{rc}, \mathcal{N}_\infty)$ defined by $P_\mu(B)=\int_E P_x(B)\mu(dx)$. Next we set for each $t\geqq 0$

(0. 10) $\mathcal{F}_t=\{B\in\mathcal{F}$; for every $\mu\in\mathfrak{M}$ there exists $B_\mu\in\mathcal{N}_t$ such that $P_\mu[B\Delta B_\mu]=0\}$[10].

Then $\mathcal{F}_{t+0}=\mathcal{F}_t$ is automatically satisfied and we may take $\mathcal{B}_t=\mathcal{F}_t$, (cf. Meyer [30]).

Definition 0.5. A Markov process $X=(X_t, \mathcal{B}_t, \zeta(\omega))$ on E with $\Delta\in E$ as the terminal point is called a *standard process* if it satisfies the following conditions;

(S. 1) $(X_t, \mathcal{B}_t)$ is strong Markov,

9) $\mathfrak{M}$=the set of all probability measures on $(E, \mathcal{B}(E))$.

10) $B\Delta B_\mu$ is the symmetric difference of B and B_μ; $B\Delta B_\mu=B\cup B_\mu-B\cap B_\mu$

(S. 2) $(X_t, \mathcal{B}_t)$ is quasi-left continuous befere ζ,

i.e., for any sequence $T_n \uparrow T$ of $\mathcal{B}_t$-Markov times, we have $P_x[X_{T_n} \to X_T, T<\zeta] = P_x[T<\zeta]$, for every $x \in E$.

(S. 3) $\overline{\mathcal{B}}_{t+0} = \mathcal{B}_t$.

Now consider a Markov process $X=(X_t, \mathcal{B}_t)$ with the terminal point Δ; $\mathcal{N}_t$, $\mathcal{N}_\infty$, $\overline{\mathcal{N}}_t$ and $\overline{\mathcal{N}}_\infty$ are defined as above.

Definition 0.6. A function $A(t, \omega)$; $(t, \omega) \in [0,\infty) \times \Omega \to A(t, \omega) \in (-\infty, \infty]$ is called a *$\mathcal{B}_t$-additive functional* if

(i) for fixed $t \in [0, \infty)$, it is $\overline{\mathcal{N}}_\infty \cap \mathcal{B}_t$-measurable and

(ii) there exists $\Omega_A \subset \Omega$,[11] $\Omega_A \in \overline{\mathcal{N}}_\infty$, $\theta_t(\Omega_A) \subset \Omega_A$ for all $t \geqq 0$, and $P_x[\Omega_A]=1$ for all $x \in E$, such that if $\omega \in \Omega_A$

1) $t \to A_t(\omega)$ is right continuous,

2) $A_0(\omega)=0$, $A_t(\omega)=A_{\zeta(\omega)}(\omega)$ for $t \geqq \zeta(\omega)$, and

3) $A_{t+s}(\omega)=A_s(\omega)+A_t(\theta_s\omega)$ for all $t, s \geqq 0$.

If we have further (for $\omega \in \Omega_A$)

4) $t \to A_t(\omega)$ is continuous,

or

5) $A_t(\omega) \geqq 0$,

then we shall call $A(t, \omega)$ *continuous or non-negative*, respectively.

Definition 0.7. A function $M_t(\omega)$; $(t, \omega) \in [0, \infty) \times \Omega \to M_t(\omega) \in [0, \infty)$ is called a *$\mathcal{B}_t$-multiplicative functional* if $A_t(\omega) = -\log M_t(\omega)$ is a $\mathcal{B}_t$-additive functional.

Given a standard process $X=(X_t, P_x)$ on E with $\Delta \in E$ as the terminal point and an $\mathcal{N}_t$-multiplicative functional M_t of X such that $E_x[M_t] \leqq 1$ for every x and $t \geqq 0$, there exists a unique (up to equivalence) standard process $\widetilde{X}=(\widetilde{X}_t, \widetilde{P}_x)$ on E with $\Delta \in E$ as the

11) Such a Ω_A is called a *defining set* for A.

terminal point such that

$$\tilde{P}_x[\tilde{X}_t \in A] = E_x[M_t;\ X_t \in A],\quad A \in \mathcal{B}(E). \tag{0.11}$$

(cf. Dynkin [6] and Kunita and T. Watanabe [28]).

Definition 0.8. $\tilde{X}$ is called the *M_t-subprocess of X.*

When $M_t(\omega) = e^{-A_t(\omega)}$, where $A_t(\omega)$ is a non-negative and continuous $\mathcal{N}_t$-additive functional, a version of the M_t-subprocess is obtained by the following method of curtailment of the life time. Let $Z(\omega)$ be a $\mathcal{B}_\infty$-measurable random variable such that $P_x[Z(\omega) > t \mid \mathcal{N}_\infty] = e^{-t}$ for every $x \in E$ and t.[12]
Put

$$\bar{\zeta}(\omega) = \begin{cases} \inf\{t;\ A_t(\omega) \geqq Z(\omega)\}, \\ \zeta(\omega),\ \text{if}\ \{\quad\} = \emptyset, \end{cases} \tag{0.12}$$

$$\bar{X}_t(\omega) = \begin{cases} X_t(\omega), & t < \bar{\zeta}(\omega), \\ \Delta, & t \geqq \bar{\zeta}(\omega). \end{cases} \tag{0.13}$$

Then the stochastic process $\{\bar{X}_t(\omega), P_x\}$ is equivalent to the M_t-subprocess $\tilde{X} = (\tilde{X}_t, \tilde{P}_x, \tilde{\zeta})$ of X.

Let $X = (X_t, \mathcal{B}_t)$, where $\mathcal{B}_t$ satisfies $\bar{\mathcal{B}}_{t+0} = \mathcal{B}_t$, be a right continuous strong Markov process such that with probability one for all P_x the left hand limits of X_t exist.

Definition 0.9. A system $(n(x, dy), A_t)$, where $n(x, dy)$ is a non-negative kernel on $E \times E$[13] and A_t is a continuous and nonnegative additive functional, is called a *Lévy system of the process X* if for every $f \in \boldsymbol{B}^+(E \times E)$ such that $f(x, x) = 0$, we have

$$E_x\Big[\sum_{s \leq t} f(X_{s-}, X_s)\Big] = E_x\Big[\int_0^t Nf(X_s)\,dA_s\Big], \tag{0.14}$$

where

12) By enlarging Ω and $\mathcal{B}_t$ if necessary, we can always assume such Z exists.

13) Let $(S, \mathcal{B})$ and $(S', \mathcal{B}')$ be two measurable spaces. $\nu(x, A)$, $x \in S$, $A \in \mathcal{B}'$ is called a ***kernel on*** $(S, \mathcal{B}) \times (S', \mathcal{B}')$ if for fixed $A \in \mathcal{B}'$, it is a $\mathcal{B}$-measurable function of x, and for fixed $x \in \mathcal{B}$, it is a measure on $(S', \mathcal{B}')$. When S and S' are topological spaces and $\mathcal{B}$ and $\mathcal{B}'$ are topological Borel fields we call it simply a kernel on $S \times S'$.

(0.15) $$Nf(x)=\int_E n(x, dy)f(x, y).$$

It is known [45] that every Hunt process with a reference measure[14] possesses a Lévy system. Suppose X possesses a Lévy system $(n(x, dy), A)$ and let D be an open set of E. Let $\tau_D=\inf\{t;\ X_t(\omega)\notin D\}$. For $B_1\subset D$ and $B_2\subset E-D$, $B_i\in\mathscr{B}(E)$ $(i=1, 2)$ and $f(x, y)=I_{B_1}(x)I_{B_2}(y)$,[15] we have from (0.14)

$$E_x[e^{-\lambda\tau_D};\ X_{\tau_D-}\in B_1,\ X_{\tau_D}\in B_2]$$
$$=E_x[\sum_{s\leq\tau_D}e^{-\lambda s}f(X_{s-},\ X_s)]=E_x\left[\int_0^{\tau_D}e^{-\lambda s}I_{B_1}(X_s)n(X_s,\ B_2)dA_s\right],$$

where λ is a positive constant. In particular,

$$E_x[e^{-\lambda\tau_D};\ X_{\tau_D-}\in B_1]=E_x\left[\int_0^{\tau_D}e^{-\lambda s}I_{B_1}(X_s)n(X_s,\ E-D)dA_s\right].$$

Therefore, if we define a kernel by

$$\Pi_D(x, B)=\frac{n(x, B)}{n(x, E-D)},$$

where $x\in D$, $B\subset E-D$, and $B\in\mathscr{B}(E)$, then we have

$$E_x[e^{-\lambda\tau_D};\ X_{\tau_D-}\in B_1,\ X_{\tau_D}\in B_2]$$
$$=E_x\left[\int_0^{\tau_D}e^{-\lambda s}\, I_{B_1}(X_s)n(X_s,\ E-D)\Pi_D(X_s,\ B_2)dA_s\right]$$
$$=E_x[e^{-\lambda\tau_D}\, I_{B_1}(X_{\tau_D-})\Pi_D(X_{\tau_D-},\ B_2)].$$

As a consequence of this formula we have

Theorem 0.2. *Suppose X possesses a Lévy system; then we have for every open set $D\subset E$ and $B\in\mathscr{B}(E)$ such that $B\subset E-D$,*

(0.16) $$P_x[X_{\tau_D}\in B\,|\,X_{\tau_D-}]=\Pi_D(X_{\tau_D-},\ B)$$

and further

(0.17) $$E_x[e^{-\lambda\tau_D}I_B(X_{\tau_D})\,|\,X_{\tau_D-}]=\Pi_D(X_{\tau_D-},\ B)E_x[e^{-\lambda\tau_D}|\,X_{\tau_D-}]$$

14) $m(dx)$ is a *reference measure for* X if for every $F\in\mathscr{B}(E)$

$$\int_E m(dx)E_x\left[\int_0^\infty e^{-t}I_F(X_t)dt\right]=0 \text{ implies } E_x\left[\int_0^\infty e^{-t}I_F(X_t)dt\right]=0$$

for every x. (This was first introduced as the condition (L) in Meyer [30]).

15) $I_B(x)$ is the indicator function of a set B.

a.s. (P_x) *on* $\{X_{\tau_D -}\in D,\ \tau_D > \infty\}$.

§0.2. Symmetric product spaces and their direct sum

Let S be a compact Hausdorff space with a countable open base (i.e. a compact metrizable space), and let $S^{(n)}(n=1, 2, \cdots)$ be the n-fold product of S with the product topology. *The symmetric n-fold product space* S^n of S is the quotient space $S^{(n)}/\sim$, where $\sim$ is the equivalence relation of the permutation; i.e., $x\sim y$, $x, y\in S^{(n)}$, if y is obtained from x by a permutation of coordinates. By the quotient topology S^n is compact. Let $S^0=\{\partial\}$, where ∂ is an extra point, and let $\boldsymbol{S}$ be the topological sum of S^n, $n=0, 1, 2, \cdots$.(16) Then $\boldsymbol{S}$ is a locally compact and non-compact Hausdorff space with a countable open base; let

$$(0.18) \qquad \hat{\boldsymbol{S}}=\boldsymbol{S}\cup\{\Delta\}$$

be the one-point compactification of $\boldsymbol{S}$. It is convenient to introduce the notation $S^{\infty}=\{\Delta\}$. Then $\hat{\boldsymbol{S}}$ is the sum of S^n, $n=0, 1, 2, \cdots, +\infty$.

Example 0.1. Assume that S consists of a single point: $S=\{a\}$. Then $S^n=\{[\overbrace{a, a, \cdots, a}^{n}]\}$, which we can identify with n. Thus $\boldsymbol{S}$ can be identified with the set $\boldsymbol{Z}^+$ of all non-negative integers, and $\hat{\boldsymbol{S}}$ with $\hat{\boldsymbol{Z}}^+=\boldsymbol{Z}^+\cup\{+\infty\}$ the compactification of $\boldsymbol{Z}^+$. More generally if S consists of k-points; $S=\{a_1, a_2, \cdots, a_k\}$, then

$$S^n=\{[\overbrace{a_1, \cdots, a_1}^{n_1}, \overbrace{a_2, \cdots, a_2}^{n_2}, \cdots, \overbrace{a_k, \cdots, a_k}^{n_k}]\ ;\ n_1+n_2+\cdots+n_k=n\},$$

which we can identify with $\{(n_1, n_2, \cdots, n_k);\ n_i\in\boldsymbol{Z}^+,\ n_1+n_2+\cdots+n_k=n\}$. Accordingly $\boldsymbol{S}=\bigcup_{n=0}^{\infty} S^n$ can be identified with $\overbrace{\boldsymbol{Z}^+\times\boldsymbol{Z}^+\times\cdots\times\boldsymbol{Z}^+}^{k}=\{(n_1, n_2, \cdots, n_k);\ n_i\in\boldsymbol{Z}^+\}$.

Let ρ be the natural mapping: $S^{(n)}\to S^n$ which maps $x\in S^{(n)}$ to

---

16) Cf. Bourbaki [3] p.35. We shall write $\boldsymbol{S}=\bigcup_{n=0}^{\infty}\boldsymbol{S}^n$.

the equivalence class ρx containing x. Then ρ extends to a mapping from the sum $\bigcup_{n=0}^{\infty} S^{(n)}$ of $S^{(n)}$ to $\boldsymbol{S}$. We shall write $\rho x=[x_1, x_2, \cdots, x_n]$ when $x=(x_1, \cdots, x_n)\in S^{(n)}$.

Now we shall define for each $m=1, 2, \cdots$, a mapping γ: $\widehat{\boldsymbol{S}}\times\widehat{\boldsymbol{S}}\times\cdots\times\widehat{\boldsymbol{S}}\to\widehat{\boldsymbol{S}}$ as follows. Let $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m\in\widehat{\boldsymbol{S}}$ then there happens just one of the following three cases;

1) $\boldsymbol{x}_i=\varDelta$ for some i,
2) $\boldsymbol{x}_i=\partial$ for every i, or
3) all $\boldsymbol{x}_i$ are different from $\varDelta$, but there is some $\boldsymbol{x}_i$ different from ∂.

Then we set

$$(0.19)\qquad \gamma(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m)=\begin{cases}\varDelta, & \text{if the case 1) happens,}\\ \partial, & \text{if the case 2) happens,}\\ \rho(x_{11}, \cdots, x_{1n_1}, x_{21}, \cdots, x_{2n_2}, \cdots, x_{m1}, \\ \qquad \cdots, x_{mn_m}), & \text{if the case 3) happens,}\end{cases}$$

where we take all $\boldsymbol{x}_i=[x_{i1}, x_{i2}, \cdots, x_{in_i}]$ such that $\boldsymbol{x}_i\neq\partial$ in the case 3).

Example 0.2. In Example 0.1, in the case $S=\{a\}$, γ is given simply by

$$\gamma(n_1, n_2, \cdots, n_m)=n_1+n_2+\cdots+n_m, \quad n_i\in\widehat{\boldsymbol{Z}}^+.$$

In the case $S=\{a_1, a_2, \cdots, a_k\}$, $\boldsymbol{n}\in\widehat{\boldsymbol{S}}$ is given by $\boldsymbol{n}=(n_1, n_2, \cdots, n_k)$ or $\boldsymbol{n}=\varDelta$ and

$$\gamma(\boldsymbol{n}^1, \boldsymbol{n}^2, \cdots, \boldsymbol{n}^m)=\begin{cases}\boldsymbol{n}^1+\boldsymbol{n}^2+\cdots+\boldsymbol{n}^m, & \text{if } \boldsymbol{n}^i\neq\varDelta \text{ for all } i,\\ \varDelta\quad , & \text{if otherwise.}\end{cases}$$

Now we shall introduce the following function spaces which are supposed to be real unless otherwise stated.

$$(0.20)\qquad \boldsymbol{B}(S)=\text{the set of all bounded Borel functions on } S,$$

$$(0.21)\qquad \boldsymbol{C}(S)=\text{the set of all continuous functions on } S,$$

$$(0.22)\qquad \boldsymbol{B}_r^*(S)=\{f\in\boldsymbol{B}(S);\ \|f\|\equiv\sup_{x\in S}|f(x)|<r\},$$

$$(0.23)\qquad \boldsymbol{C}_r^*(S)=\boldsymbol{C}(S)\cap\boldsymbol{B}_r^*(S).$$

We shall denote by "——" the closure with respect to the norm, and so

(0.24) $\overline{\boldsymbol{B}_r^*}(S)=\{f\in \boldsymbol{B}(S);\ \|f\|\leq r\}$,

(0.25) $\overline{\boldsymbol{C}_r^*}(S)=\{f\in \boldsymbol{C}(S);\ \|f\|\leq r\}$.

For $r=1$, we shall omit the subscript r;

(0.26) $\boldsymbol{B}^*(S)=\boldsymbol{B}_1^*(S)=\{f\in \boldsymbol{B}(S);\ \|f\|<1\}$,

(0.27) $\boldsymbol{C}^*(S)=\{f\in \boldsymbol{C}(S);\ \|f\|<1\}$.

$\boldsymbol{B}(S^n)$, $\boldsymbol{C}(S^n)$, $\boldsymbol{B}(\boldsymbol{S})$, $\boldsymbol{C}(\boldsymbol{S})$, $\boldsymbol{B}(\widehat{\boldsymbol{S}})$ and $\boldsymbol{C}(\widehat{\boldsymbol{S}})$ are defined similarly. *When we consider $f\in \boldsymbol{B}(\boldsymbol{S})$ as a function on $\widehat{\boldsymbol{S}}$ we shall always put $f(\Delta)=0$.* The supremum norm of $\boldsymbol{B}(S^n)$(resp. $\boldsymbol{B}(\boldsymbol{S})$, $\boldsymbol{B}(\widehat{\boldsymbol{S}})$) is denoted as $\|\ \|_{S^n}$, (resp. $\|\ \|_{\boldsymbol{S}}$, $\|\ \|_{\widehat{\boldsymbol{S}}}$). Further we shall introduce

(0.28) $\boldsymbol{B}_0(\boldsymbol{S})=\{f\in \boldsymbol{B}(\boldsymbol{S});\ \lim_{\boldsymbol{x}\to\Delta} f(\boldsymbol{x})=0\}$, and

(0.29) $\boldsymbol{C}_0(\boldsymbol{S})=\boldsymbol{C}(\boldsymbol{S})\cap \boldsymbol{B}_0(\boldsymbol{S})$.

The set of all (not necessarily bounded) Borel functions are denoted as $\mathfrak{B}(S)$, $\mathfrak{B}(S^n)$, $\cdots$etc. The subclass of each function space introduced above formed of all non-negative elements is denoted by "+", *e.g.*, $\boldsymbol{B}(\boldsymbol{S})^+$, $\boldsymbol{C}(\boldsymbol{S})^+$, $\boldsymbol{B}^*(S)^+$, $\cdots$etc. For $f\in \boldsymbol{B}(\boldsymbol{S})$ or $\boldsymbol{B}(\widehat{\boldsymbol{S}})$, the restriction of f on S^n is a function in $\boldsymbol{B}(S^n)$, which we shall denote as $f|_{S^n}$. In particular $f|_S$ is the restriction f on S.

Next we shall define serveral operations on functions which will play an important role in the future discussions. First of all, "$\wedge$" is a mapping

$$\wedge:\ \overline{\boldsymbol{B}^*}(S)\ni f\to \widehat{f}\in \boldsymbol{B}(\widehat{\boldsymbol{S}})$$

defined by

(0.30)
$$\widehat{f}(\boldsymbol{x})=\begin{cases}1, & \text{if } \boldsymbol{x}=\partial,\\ f(x_1)f(x_2)\cdots f(x_n), & \text{if } \boldsymbol{x}=[x_1, x_2, \cdots, x_n],\\ 0, & \text{if } \boldsymbol{x}=\Delta.\end{cases}$$

Since $f(x_1)f(x_2)\cdots f(x_n)$ is invariant under the permutation of $x_1, x_2, \cdots, x_n$, it is well defined. Clearly "$\wedge$" maps $\boldsymbol{B}^*(S)$ into $\boldsymbol{B}_0(\boldsymbol{S})$ and $\boldsymbol{C}^*(S)$ into $\boldsymbol{C}_0(\boldsymbol{S})$. Next for $f\in\overline{\boldsymbol{B}^*}(S)$ and $g\in \boldsymbol{B}(S)$ define a function $\langle f|g\rangle\in\mathfrak{B}(\widehat{S})$ by

$$(0.31)\qquad \langle f|g\rangle(\boldsymbol{x})=\begin{cases}0, & \text{if } \boldsymbol{x}=\partial,\\ g(x), & \text{if } \boldsymbol{x}=x\in S,\\ \sum_{k=1}^{n} g(x_k)\prod_{\substack{i=1\\ i\neq k}}^{n} f(x_i), & \text{if } \boldsymbol{x}=[x_1, x_2, \cdots, x_n]\in S^n,\\ 0, & \text{if } \boldsymbol{x}=\Delta.\end{cases}$$

Clearly $\langle f|g\rangle\in \boldsymbol{B}_0(S)$ (resp. $\boldsymbol{C}_0(S)$) provided $f\in\boldsymbol{B}^*(S)$ (resp. $f\in\boldsymbol{C}^*(S)$ and $g\in\boldsymbol{C}(S)$). Further we shall introduce the following operator $\vee$ from $f\in\boldsymbol{B}(S)$ to $\check{f}\in\mathfrak{B}(\widehat{S})$ defined by

$$(0.32)\qquad \check{f}(\boldsymbol{x})=\langle 1|f\rangle(\boldsymbol{x})=\begin{cases}0, & \text{if } \boldsymbol{x}=\partial,\\ f(x_1)+f(x_2)+\cdots+f(x_n), & \\ \qquad \text{if } \boldsymbol{x}=[x_1, x_2, \cdots, x_n], & \\ 0, & \text{if } \boldsymbol{x}=\Delta.\end{cases}$$

Lemma 0.1. *For every* $0<r<1$, *there exist positive constants* a_r, b_r, c_r, d_r *and* e_r *such that*

$$(0.33)\qquad \|\hat{f}-\hat{g}\|_S\leq a_r\|f-g\| \quad \textit{for all } f, g\in \boldsymbol{B}_r^*(S),$$

$$(0.34)\qquad \|<f|u>-<g|v>\|_S\leq b_r\|u\|\cdot\|f-g\|+c_r\|u-v\|$$
$$\textit{for all } f, g\in\boldsymbol{B}_r^*(S) \textit{ and } u, v\in\boldsymbol{B}(S),$$

$$(0.35)\qquad \|\frac{1}{t}(\hat{g}-\hat{f})-<f|h>\|_S\leq d_r\|\frac{1}{t}(g-f)-h\| + e_r\|h\|\|f-g\|$$
$$\textit{for all } f, g\in\boldsymbol{B}_r^*(S),\ h\in\boldsymbol{B}(S) \textit{ and } t>0.$$

In particular, we have from (0.35) *that for every* $f\in\boldsymbol{B}^*(S)$ *and* $h\in\boldsymbol{B}(S)$

$$(0.36)\qquad \lim_{\varepsilon\to 0}\|\frac{1}{\varepsilon}\{(\widehat{f+\varepsilon h})-\hat{f}\}-<f|h>\|_S=0.$$

Proof. For $\boldsymbol{x}=[x_1, x_2, \cdots, x_n]\in S^n$ we have

$$(0.37)\qquad \hat{f}(\boldsymbol{x})-\hat{g}(\boldsymbol{x})=\prod_{i=1}^{n} f(x_i)-\prod_{i=1}^{n} g(x_i)$$
$$=\sum_{k=1}^{n}(f(x_k)-g(x_k))f(x_1)\cdots f(x_{k-1})g(x_{k+1})\cdots g(x_n).$$

Hence if $f, g \in \boldsymbol{B}_r^*(S)$ $(r<1)$, we have

$$\|\widehat{f}-\widehat{g}\|_S \leq \sup_n (nr^{n-1})\|f-g\|,$$

which proves (0.33). Next,

$$\langle f|u\rangle(\boldsymbol{x})-\langle g|v\rangle(\boldsymbol{x})=\sum_{k=1}^{n}\{u(x_k)\prod_{i\neq k} f(x_i)-v(x_k)\prod_{i\neq k} g(x_i)\}$$

$$=\sum_{k=1}^{n}(u(x_k)-v(x_k))\prod_{i\neq k} g(x_i)+\sum_{k=1}^{n}u(x_k)(\prod_{i\neq k} f(x_i)-\prod_{i\neq k} g(x_i)),$$

and by the same way as above, we have

$$|\prod_{i\neq k} f(x_i)-\prod_{i\neq k} g(x_i)|\leq(n-1)r^{n-2}\|f-g\|,$$

provided $f, g\in \boldsymbol{B}_r^*(S)$. Therefore

$$\|\langle f|u\rangle-\langle g|v\rangle\|_S\leq(\sup_n nr^{n-1})\|u-v\| + (\sup_n n(n-1)r^{n-2})\|u\|\cdot\|f-g\|$$

which proves (0.34). Finally,

$$\langle f|h\rangle(\boldsymbol{x})=\sum_{k=1}^{n} h(x_k)\prod_{i\neq k} f(x_i)$$

$$=\sum_{k=1}^{n} h(x_k)f(x_1)\cdots f(x_{k-1})g(x_{k+1})\cdots g(x_n)$$

$$+\sum_{k=1}^{n} h(x_k)f(x_1)\cdots f(x_{k-1})\{f(x_{k+1})\cdots f(x_n)-g(x_{k+1})\cdots g(x_n)\}$$

and hence by (0.37)

$$\frac{1}{t}(\widehat{g}-\widehat{f})(\boldsymbol{x})-\langle f|h\rangle(\boldsymbol{x})$$

$$=\sum_{k=1}^{n}\left\{\frac{1}{t}(g(x_k)-f(x_k))-h(x_k)\right\}f(x_1)\cdots f(x_{k-1})g(x_{k+1})\cdots g(x_n)$$

$$+\sum_{k=1}^{n} h(x_k)f(x_1)\cdots f(x_{k-1})\{f(x_{k+1})\cdots f(x_n)-g(x_{k+1})\cdots g(x_n)\}.$$

Since

$$|f(x_{k+1})\cdots f(x_n)-g(x_{k+1})\cdots g(x_n)|\leq(n-k)r^{n-k-1}\|f-g\|,$$

we have

$$\|\frac{1}{t}(\widehat{g}-\widehat{h})-\langle f|h\rangle\|_S\leq(\sup_n nr^{n-1})\|\frac{1}{t}(g-f)-h\|$$

$$+\left(\sup_n \frac{n(n-1)}{2} r^{n-2}\right)\|h\|\|f-g\|,$$

which proves (0.35).

Lemma 0.2. (i) *The linear hull of the subset* $\{\hat{f}: f \in \boldsymbol{C}^*(S)^+\}$[17] *of* $\boldsymbol{C}_0(S)$ *is dense in* $\boldsymbol{C}_0(S)$.

(ii) *For every* $T>0$, *the linear hull of the subset*

$$\{\hat{f}; f(s, x)=g(s)h(x)\in \boldsymbol{C}^*([0, T]\times S)^+\}^{(18)}$$

of $\boldsymbol{C}_0([0, T]\times S)$ *is dense in* $\boldsymbol{C}_0([0, T]\times S)$. *In particular, the linear hull of* $\{\hat{f}; f\in \boldsymbol{C}^*([0, T]\times S)^+\}$ *is dense in* $\boldsymbol{C}_0([0, T]\times S)$.

Proof. First of all, we note that the linear hull is given by $\{\sum_{i=1}^{n} c_i\hat{f}_i;\ f_i\in \boldsymbol{C}^*(S^+), c_i$: real constants$\}$. By the Stone-Weierstrass theorem, the linear hull of $\{f_1(x_1)f_2(x_2)\cdots f_n(x_n);\ f_i\in \boldsymbol{C}(S)^+\}$ is dense *in* $\boldsymbol{C}(S^{(n)})$, and hence the linear hull of $\{\sum_{\pi}\prod_{i=1}^{n} f_{\pi(i)}(x_i);\ f_i\in \boldsymbol{C}^*(S)^+\}$ is dense in $\boldsymbol{C}(S^n)$.[19] But by a combinatorial lemma,[20] we have

$$\sum_{\pi}\prod_{i=1}^{n} f_{\pi(i)}(x_i)=\prod_{i=1}^{n}\Big(\sum_{k=1}^{n} f_k(x_i)\Big)-\sum_{(k_1,\cdots,k_{n-1})}\prod_{i=1}^{n}\Big(\sum_{q=1}^{n-1} f_{k_q}(x_i)\Big)$$
$$+\sum_{(k_1,\cdots,k_{n-2})}\prod_{i=1}^{n}\Big(\sum_{q=1}^{n-2} f_{k_q}(x_i)\Big)-\cdots+(-1)^{n-1}\sum_{k}\prod_{i=1}^{n} f_k(x_i),$$

where $\sum_{(k_1,k_2,\cdots,k_r)}$ denotes the sum over all $(k_1, k_2, \cdots, k_r)$ such that $1\leq k_i\leq n$ and all k_i are different. This implies $\sum_{\pi}\prod_{i=1}^{n} f_{\pi(i)}(x_i)$ belongs to the linear hull of $\{f(x_1)\cdots f(x_n);\ f\in \boldsymbol{C}^*(S)^+\}=\{\hat{f}|_{S^n};\ f\in \boldsymbol{C}^*(S)^+\}$ and hence the linear hull of $\{\hat{f}|_{S^n};\ f\in \boldsymbol{C}^*(S)^+\}$ is dense in $\boldsymbol{C}(S^n)$.

Now in order to prove the linear hull of $\{\hat{f};\ f\in \boldsymbol{C}^*(S)^+\}$ is dense in $\boldsymbol{C}_0(S)$, it is sufficient to show that any continuous linear func-

---

17) $\boldsymbol{C}^*(S)^+=\{f\in \boldsymbol{C}^*(S), f\geq 0\}=\{f:$ continuous on S, $0\leq f<1\}$.

18) $\boldsymbol{C}([0, T]\times S)=\{f=f(t, x);$ continuous on $[0, T]\times S\}$, $\boldsymbol{C}^*([0, T]\times S)^+=\{f=f(t, x);$ continuous on $[0, T]\times S$, $0\leq f<1\}$ and $\hat{f}=\hat{f}(t, x)$ is defined by (0.30) for each fixed t.

19) $\sum_{\pi}$ denotes the sum over all permutations π on $(1, 2, \cdots, n)$.

20) Cf. Ryser [38].

tional μ on $\boldsymbol{C}_0(S)$, (i.e., any signed Radon measure $\mu(dx)$ on $\boldsymbol{S}$ with finite total variation) such that $\mu(\widehat{f})=\int_S \widehat{f}(x)\mu(dx)=0$ for every $f \in \boldsymbol{C}^*(S)^+$ is identically zero. Suppose, therefore, that $\mu(\widehat{f})=0$ for every $f\in \boldsymbol{C}^*(S)^+$, then for every $\lambda(0\leqq\lambda\leqq 1)$, $\lambda f\in \boldsymbol{C}^*(S)^+$, and hence $\mu(\widehat{\lambda f})=\sum_{n=0}^{\infty}\lambda^n\int_{S^n}\widehat{f}(x)\mu\ (dx)=0$. Thus we have $\int_{S^n}\widehat{f}(x)\mu(dx)=0$ for every $f\in \boldsymbol{C}^*(S)^+$ $(n=0, 1, 2, \cdots)$. But we have shown above that the linear hull of $\{\widehat{f}|_{S^n};\ f\in \boldsymbol{C}^*(S)^+\}$ is dense in $\boldsymbol{C}(S^n)$, and hence $\mu|_{S^n}=0$ for every $n=0, 1, 2, \cdots$. This proves $\mu=0$, and the proof of (i) is complete. (ii) can be proved in a similar way.

Lemma 0.3. (i) *Let*

(0.38) $\mathfrak{R}=$ *the set of all signed Radon measures* μ *on* $\boldsymbol{S}$ *with finite total variations.*

Then for $\nu_1, \nu_2, \cdots, \nu_k\in\mathfrak{R}$, *there exists one and only one* $\mu\in\mathfrak{R}$ *such that*

(0.39) $$\mu(\widehat{f})=\prod_{i=1}^{k}\nu_i(\widehat{f}),\ f\in \boldsymbol{C}^*(S)^+.$$

We shall denote this μ *as*

(0.40) $$\mu=\nu_1*\nu_2*\cdots*\nu_k,$$

then

(0.41) $$|\mu|=|\nu_1|*|\nu_2|*\cdots*|\nu_k|^{(21)},$$ *and*

(0.42) $$\mu(\boldsymbol{S})=\prod_{i=1}^{k}\nu_i(\boldsymbol{S}).$$

Hence, in particular, μ *is positive (resp. a probability measure) if all* ν_i *are positive (resp. probability measures).*

(ii) *Let* $\nu_0(dt, dx)$ *be a signed Radon measure on* $[0, \infty)\times \boldsymbol{S}$ *with finite total variation and* $\nu_1(t, \cdot), \cdots, \nu_k(t, \cdot)\in\mathfrak{R}$ *such that, for every* $E\in\mathcal{B}(\boldsymbol{S})$, $\nu_i(t, E)$ *is a bounded Borel measurable function in* t *then there exists one and only one* $\mu(dt, dx)$, *a signed measure on* $[0, \infty)\times \boldsymbol{S}$ *with finite total variation such that*

21) $|\mu|=\mu^+ + \mu^-$, where $\mu=\mu^+ - \mu^-$ is the Jordan decomposition of $\mu\in\mathfrak{R}$.

(0.43) $$\int_{[0,\infty)\times\boldsymbol{S}} \widehat{f}(s, \boldsymbol{x})\mu(ds, d\boldsymbol{x}) = \int_{[0,\infty)\times\boldsymbol{S}} \widehat{f}(s, \boldsymbol{x})\nu_0(ds, d\boldsymbol{x}) \prod_{i=1}^{k} \nu_i(s, \widehat{f}),$$
$$f \in \boldsymbol{C}^*([0, \infty) \times S)^+,$$

where $\nu_i(s, \widehat{f}) = \int_{\boldsymbol{S}} \nu_i(s, d\boldsymbol{x})\widehat{f}(s, \boldsymbol{x})$. *We shall denote this* μ *as*

(0.44) $$\mu = \nu_0 \otimes \nu_1 * \nu_2 * \cdots * \nu_k$$

then

(0.45) $$|\mu| = |\nu_0| \otimes |\nu_1| * |\nu_2| * \cdots * |\nu_k|$$

and

(0.46) $$\mu([0, \infty) \times \boldsymbol{S}) = \iint_{[0,\infty)\times\boldsymbol{S}} \nu_0(ds, d\boldsymbol{x}) \prod_{i=1}^{k} \nu_i(s, \boldsymbol{S}).$$

Hence μ *is positive (resp. a probability measure) if* $\nu_0, \nu_1, \cdots, \nu_k$ *are positive (resp. probability measures).*

Proof. It is sufficient to prove the case $k=2$; then

$$\nu_1(\widehat{f})\nu_2(\widehat{f}) = \sum_{n,m=0}^{\infty} \iint_{S^n \times S^m} f(x_1)\cdots f(x_n) f(y_1)\cdots f(y_m)\nu_1(d\boldsymbol{x})\nu_2(d\boldsymbol{y}).$$

Let γ be the mapping $\boldsymbol{S} \times \boldsymbol{S} \to \boldsymbol{S}$ defined in (0.19) then $\mu \in \mathcal{R}$ defined by

$$\mu|_{S^n} = \sum_{n=k+j} (\nu_1|_{S^k} \times \nu_2|_{S^j}) \circ \gamma^{-1}, \qquad k, j = 0, 1, 2, \cdots,$$

clearly satisfies $\mu(\widehat{f}) = \nu_1(\widehat{f}) \cdot \nu_2(\widehat{f})$. The uniqueness follows from Lemma 0.2. (0.41) and (0.42) follow from the definition and the following property of product measures $|\lambda \times \mu| = |\lambda| \times |\mu|$. (ii) can be proved in a similar way.

Example 0.3. In Example 0.1, if $S = \{a\}$, then $\boldsymbol{S} = \boldsymbol{Z}^+$ and hence $\mathcal{R} = \{(a_n)_{n=0}^{\infty}$ such that $\Sigma |a_n| < \infty\}$. $f \in \boldsymbol{C}^*(S)^+$ is determined by a constant λ such that $0 \leq \lambda < 1$ and $\widehat{f}(i) = \lambda^i$. If $\mu = (a_n)_{n=0}^{\infty}$ then $\mu(\widehat{f}) = \sum_{n=0}^{\infty} a_n \lambda^n$, which is nothing but the generating function of μ. If $\nu = (b_n)_{n=0}^{\infty}$ then $\mu * \nu = (c_n)_{n=0}^{\infty}$, where $c_n = \sum_{k+l=n} a_k b_l$. We remark also $\check{f}(i) = i\lambda$.

If $S=\{a_1, a_2, \cdots, a_k\}$, then $\boldsymbol{S}=\{\boldsymbol{n}=(n_1, n_2, \cdots, n_k);\ n_i\in \boldsymbol{Z}^+\}$ and $f\in \boldsymbol{C}^*(S)^+$ is determined by constants $(\lambda_1, \cdots, \lambda_k)$, $0\leqq\lambda_i<1$ by the relation $\lambda_i=f(\boldsymbol{e}_i)$ where $\boldsymbol{e}_i=(0, \cdots, 0, \overset{i-th}{1}, 0, \cdots, 0)$. Then $\widehat{f}(\boldsymbol{n})=\lambda_1^{n_1}\lambda_2^{n_2}\cdots\lambda_k^{n_k}$. $\mathcal{R}=\{(a_{n_1 n_2\cdots n_k})$ such that $\sum_{n_1\cdots n_k}|a_{n_1\cdots n_k}|<\infty\}$ and $\mu(\widehat{f})=\Sigma a_{n_1 n_2\cdots n_k}\lambda_1^{n_1}\lambda_2^{n_2}\cdots\lambda_k^{n_k}$. If $\nu=(b_{n_1, n_2\cdots n_k})\in\mathcal{R}$, then $\mu*\nu=(c_{n_1, n_2\cdots n_k})$, where $c_{n_1 n_2\cdots n_k}=\sum_{\substack{n_i=j_i+l_i \\ i=1,2,\cdots k}} a_{j_1\cdots j_k}\, b_{l_1 l_2\cdots l_k}$. Finally, $\check{f}(\boldsymbol{n})=n_1\lambda_1+n_2\lambda_2+\cdots+n_k\lambda_k$.

I. Branching Markov processes

A branching process is a random motion of particles each producing new particles of the same character or dying out and new born particles will continue the motion independently each other. Let particles move on a topological space S then if, at time t, there exists n particles, they define a point $\boldsymbol{X}_t\in S^n$ where S^n is the n-fold symmetric direct product of S defined in §0.2. Thus, we have a stochastic process $\boldsymbol{X}_t$, whose state space is $\widehat{\boldsymbol{S}}=\boldsymbol{S}\cup\{\varDelta\}$, where $\boldsymbol{S}$ is the sum of $S^n, n=0, 1, 2,\cdots$.(22) If the motion of a particle is Markovian, then the process $\boldsymbol{X}_t$ will be a Markov process. It is, therefore, natural to define a branching process as a Markov process $\boldsymbol{X}$ on $\widehat{\boldsymbol{S}}$ with the independence property of particles. The independence property of particles can be formulated in many ways; so we will adopt one of them as the definition in §1.1 and give several equivalent formulations in §1.2. Similar formulations of a branching process were given by several authors, especially by Moyal [35] and Skorohod [43].(23) Finally, we shall show that under certain general conditions every branching Markov process $\boldsymbol{X}$ is uniquely determined by a Markov process X^0 on S, called the *non-branching part* of $\boldsymbol{X}$ (Definition 1.2), and a stochastic kernel $\pi(x, dy)$ on $S\times\widehat{\boldsymbol{S}}$, called the *branching law* of $\boldsymbol{X}$ (Definition 1.3); X^0 describes the behavior of

22) $S^{(0)}=\{\partial\}$. ∂ is the state that all particles died out and $\varDelta$ is the state of explosion.

23) Cf. also Harris [8] (Chapter III), Mullikin [36] and Silverstein [42] for different formulations of branching process.

a particle of X before its first splitting, and $\pi(x, dy)$ describes the law of the splitting. This process X will be called the (X^0, π) *-branching Markov process.* The construction of a (X^0, π)-branching Markov process for given X^0 and π will be discussed in Chapter III and Chapter IV.

§1.1 Definitions

Let S be a compact Hausdorff space with a countable open base, and define the symmetric product spaces S^n, $n=0, 1, 2, \cdots$, their direct sum $\boldsymbol{S}=\bigcup_{n=0}^{\infty} S^n$ and its compactification $\widehat{\boldsymbol{S}}=\boldsymbol{S}\cup\{\Delta\}$ as in §0. 2. Let $\boldsymbol{X}=(\Omega, \mathcal{B}_t, \boldsymbol{P}_{\boldsymbol{x}}, \boldsymbol{x}\in\widehat{\boldsymbol{S}}, \boldsymbol{X}_t, \theta_t)$ be a right-continuous Markov process on $\widehat{\boldsymbol{S}}$, and let $\boldsymbol{T}_t$ be the semi-group on $\boldsymbol{B}(\widehat{\boldsymbol{S}})$ induced by $\boldsymbol{X}$, i.e.,

$$(1.1) \qquad \boldsymbol{T}_t f(\boldsymbol{x})=\boldsymbol{E}_{\boldsymbol{x}}[f(\boldsymbol{X}_t)], \ f\in\boldsymbol{B}(\widehat{\boldsymbol{S}}).$$

Definition 1.1. A Markov process $\boldsymbol{X}$ on $\widehat{\boldsymbol{S}}$ is called a *branching Markov process* if it satisfies

$$(1.2) \qquad \boldsymbol{T}_t \widehat{f}(\boldsymbol{x})=\widehat{(\boldsymbol{T}_t\widehat{f})|_S}(\boldsymbol{x}), \ \boldsymbol{x}\in\widehat{\boldsymbol{S}},$$

for every $f\in\boldsymbol{B}^*(S)$.[24]

By taking $f\in\boldsymbol{C}^*(S)$, $0<f<1$, we have $\boldsymbol{T}_t\widehat{f}(\Delta)=\widehat{\boldsymbol{T}_t\widehat{f}}(\Delta)=0$, but $\widehat{f}(\boldsymbol{x})>0$ for all $\boldsymbol{x}\in\boldsymbol{S}$. Hence, we have $\boldsymbol{P}_\Delta[\boldsymbol{X}_t=\Delta]=1$ for every $t\geqq 0$ and, by the right continuity of $\boldsymbol{X}_t$,

$$\boldsymbol{P}_\Delta[\boldsymbol{X}_t=\Delta, \text{ for all } t\geqq 0]=1.$$

Quite similarly we have

$$\boldsymbol{P}_\partial[\boldsymbol{X}_t=\partial, \text{ for all } t\geqq 0]=1.$$

Now suppose $\boldsymbol{X}$ is strong Markov such that $\overline{\mathcal{B}}_{t+0}=\mathcal{B}_t$ then, since the hitting time $e_\Delta(e_\partial)$ for Δ (resp. ∂), where

$$e_\Delta(\text{resp. } e_\partial)=\inf\{t;\ \boldsymbol{X}_t=\Delta \ (\text{resp. } \partial)\},$$

24) Clearly (1.2) is true for every $f\in\boldsymbol{B}^*(S)$ if it is true for every $f\in\boldsymbol{C}^*(S)$. Then (1.2) is true for $f\in\overline{\boldsymbol{B}^*(S)}$.

is a Markov time (cf. Remark 0.1 (iv)),

$$\begin{aligned} &\boldsymbol{P}_x[e_\Delta=+\infty \text{ or for all } e_\Delta\leqq t<\infty,\ \boldsymbol{X}_t=\Delta] \\ &=\boldsymbol{P}_x[e_\Delta=+\infty]+\boldsymbol{E}_x[\boldsymbol{P}_{\boldsymbol{X}_{e_\Delta}}[\boldsymbol{X}_t=\Delta \text{ for all } t\geqq 0]\,;\ e_\Delta<+\infty]] \\ &=\boldsymbol{P}_x[e_\Delta=+\infty]+\boldsymbol{E}_x[\boldsymbol{P}_\Delta[\boldsymbol{X}_t=\Delta \text{ for all } t\geqq 0]\,;\ e_\Delta<+\infty]=1. \end{aligned}$$

Thus, we have the following

Theorem 1.1. *If a branching Markov process $\boldsymbol{X}=(\boldsymbol{X}_t, \mathcal{B}_t)$ is strong Markov such that $\overline{\mathcal{B}}_{t+0}=\mathcal{B}_t$, then Δ and ∂ are traps.*

Example 1.1. Consider the simplest case when $S=\{a\}$ (cf. Example 0.1 and 0.3), then $\boldsymbol{S}\sim\boldsymbol{Z}^+=\{0, 1, 2, \cdots\}$ and $\widehat{\boldsymbol{S}}\sim\widehat{\boldsymbol{Z}}^+=\{0, 1, 2, \cdots, +\infty\}$. Every $f\in \boldsymbol{C}^*(S)^+$ is given by a real number λ, $0\leqq\lambda<1$, and $f(i)=\lambda^i$. Hence

$$\boldsymbol{T}_t\widehat{f}(i)=\sum_{j=0}^{\infty}P_{ij}(t)\lambda^j,$$

where $\{P_{ij}(t)\}$ is the transition matrix. Therefore a Markov chain on $\widehat{\boldsymbol{Z}}^+$ is a branching Markov process if and only if its transition matrix $\{P_{ij}(t)\}$ satisfies

$$(1.3)\qquad \begin{cases} \sum_{j=0}^{\infty}P_{ij}(t)\lambda^j=\left(\sum_{j=0}^{\infty}P_{1j}(t)\lambda^j\right)^i, & 0\leqq\lambda<1,\ i=0, 1, 2, \cdots, \\ P_{+\infty,+\infty}(t)\equiv 1. \end{cases}$$

This is equivalent to

$$(1.4)\qquad P_{ij}(t)=\sum\prod_{m=1}^{i}P_{1,r_m}(t),\quad t\geqq 0,\ i, j\in\boldsymbol{Z}^+,$$

where the summation is taken over all $(r_1, r_2, \cdots, r_i)$ such that $r_m\in\boldsymbol{Z}^+$ and $r_1+r_2+\cdots+r_i=j$. (1.4) was adopted by Kolmogorov-Dmitriev [24] as the definition of the single-type branching process with continuous time.

In the case $S=\{a_1, a_2, \cdots, a_k\}$, $\boldsymbol{S}\sim(\boldsymbol{Z}^+)^k\equiv\{\boldsymbol{n}=(n_1, n_2, \cdots, n_k)\,;\ n_i\in\boldsymbol{Z}^+\}$ and $\widehat{\boldsymbol{S}}\sim(\boldsymbol{Z}^+)^k\cup\{+\infty\}$. Then, quite similarly, we see that a Markov chain $\boldsymbol{X}_t$ on $(\boldsymbol{Z}^+)^k\cup\{+\infty\}$ is a branching Markov process if and only if its transition matrix $\{P_{n,m}(t)\}$ satisfies

$$(1.5)\qquad \begin{cases} \sum_{m} P_{n,m}(t)\lambda_1^{m_1}\lambda_2^{m_2}\cdots\lambda_k^{m_k} = \Pi_{i=1}^{k}(\sum_{m} P_{e_i,m}(t)\lambda_1^{m_1}\lambda_2^{m_2}\cdots\lambda_k^{m_k})^{n_i} \\ P_{+\infty,+\infty}(t)\equiv 1, \end{cases}$$

where $m=(m_1, m_2, \cdots, m_k)$, $n=(n_1, n_2, \cdots, n_k)$, $0 \leqq \lambda_i < 1$ and $e_i=(0, 0, \cdots, \overset{i-th}{1}, 0, \cdots, 0)$, $(i=1, 2, \cdots, k)$.

From now on, unless otherwise stated, all branching Markov processes $X=(X_t, \mathcal{B}_t)$ *we shall consider are supposed to be strong Markov such that* $\overline{\mathcal{B}_{t+0}}=\mathcal{B}_t$.[25]

According to the intuitive meaning explained at the beginning of this chapter, we shall define

$$(1.6)\qquad \xi_t(\omega)=n, \text{ if } X_t(\omega)\in S^n,\ n=0, 1, 2, \cdots, +\infty,^{(26)}$$

and call it the *number of particles at time t*. The hitting time e_Δ for Δ and e_∂ for ∂ are defined above and we shall call them the *explosion time* and the *extinction time*, respectively. Further, we shall define

$$(1.7)\qquad \tau(\omega)=\inf\{t;\ \xi_t(\omega)\neq\xi_0(\omega)\},$$

and define $\tau_0, \tau_1, \tau_2, \cdots$, inductively by

$$(1.8)\qquad \tau_0(\omega)=0,\ \tau_1(\omega)=\tau(\omega) \text{ and}$$
$$\tau_n(\omega)=\tau_{n-1}(\omega)+\tau(\theta_{\tau_{n-1}(\omega)}\omega).$$

We shall set also

$$(1.9)\qquad \tau_\infty(\omega)=\lim_{n\to\infty}\tau_n(\omega).$$

$\tau_1, \tau_2, \cdots, \tau_\infty$ are all Markov times and τ_n is called the *n-th splitting* (or *branching*) *time* of the process X.

It should be remarked that many important quantities for the process X can be expressed in the form $T_t\widehat{f}$, $f\in\overline{B^*(S)}$. For ex-

25) Also we shall take Δ as the terminal point of X; so e_Δ is identified with the lifetime ζ.

26) $S^\infty=\{\Delta\}$.

ample,

$$\boldsymbol{P}_x[e_\Delta > t] = \boldsymbol{T}_t\widehat{f}(\boldsymbol{x}), \quad \text{for } f\equiv 1, \tag{1.10}$$

$$\boldsymbol{P}_x[e_\partial \leq t] = \boldsymbol{T}_t\widehat{f}(\boldsymbol{x}), \quad \text{for } f\equiv 0, \tag{1.11}$$

and

$$\boldsymbol{E}_x[\lambda^{\xi_t}] = \boldsymbol{T}_t\check{f}(\boldsymbol{x}), \text{ for } f\equiv\lambda,\ 0\leq\lambda<1. \tag{1.12}$$

For $f\in \boldsymbol{B}^+(S)$, define $\xi_t^f(\omega)$ by

$$\xi_t^f(\omega) = \check{f}(\boldsymbol{X}_t(\omega)). \tag{1.13}$$

If $f=I_D$, where I_D is the indicator function of $D\in\mathcal{B}(S)$, we write ξ_t^D instead of $\xi_t^{I_D}$. ξ_t^D $(t<e_\Delta)$ is the number of particles in D at time t.

Remark 1.1. For a certain problem, it happens that S contains a point ∇ such that if we set $T=\{\partial, \nabla, [\nabla,\nabla], [\nabla,\nabla,\nabla], \cdots\}$, then $\boldsymbol{P}_x[\boldsymbol{X}_t\in T$ for all $t\geqq s$ if $\boldsymbol{X}_s\in T]=1$, $\boldsymbol{x}\in S$, and it may be natural to call $\xi_t'(\omega)=\xi_t^{S-\{\nabla\}}\{\omega\}$ as the number of particles. Then the extinction time is the first hitting time e_T for the set T. (Cf. Example 3.4 (C) of Chapter III, where $S=D\cup\{\nabla\}$ is the one-point compactification of a bounded domain D in R^N).

Now let $\boldsymbol{X}=(\Omega, \mathcal{B}_t, \boldsymbol{X}_t, \boldsymbol{P}_x, \boldsymbol{x}\in\widehat{S}, \theta_t)$ be a branching Markov process, and for each $n=1,2,\cdots$ define a new Markov process X_n^0 on $S^n\cup\{\Delta\}$ in the following way:

$$X_n^0=(\Omega^0, \mathcal{B}_t^0, X_t^0, P_x^0, \boldsymbol{x}\in S^n\cup\{\Delta\}),$$

where $\Omega^0=\Omega$, $\mathcal{B}_t^0=\mathcal{B}_t$, $P_x^0=\boldsymbol{P}_x$, $\boldsymbol{x}\in S^n\cup\{\Delta\}$, and

$$X_t^0=\begin{cases} X_t(\omega), & \text{if } t<\tau(\omega),\\ \Delta, & \text{if } t\geqq\tau(\omega).\end{cases}$$

The point Δ is considered as the terminal point of X_n^0, and so $\tau(\omega)$ is identified with its life time $\zeta^0(\omega)$.[27]

27) To be precise, $P_x^0[\tau(\omega)=\zeta^0(\omega)]=1$, $\boldsymbol{x}\in S^n$, but $P_\Delta^0[\tau(\omega)=+\infty]=1$ and $P_\Delta^0[\zeta^0(\omega)=0]=1$.

Definition 1.2. X_n^0 is called the *non-branching part of* $\boldsymbol{X}$ *on* S^n. $X^0 \equiv X_1^0$ is called simply the *non-branching part of* $\boldsymbol{X}$.

Thus, the non-branching part of $\boldsymbol{X}$ is a strong Markov process X^0 on $S \cup \{\Delta\}$ with Δ as its terminal point, which will describe the behavior of a particle of $\boldsymbol{X}$ before its first splitting.

Next we shall consider the following

Assumption 1.1. $\boldsymbol{X}_{\tau-} = \lim_{n\to\infty} \boldsymbol{X}_{\tau-\frac{1}{n}}$ exists almost surely on $\{\tau < \infty\}$, and there exists a stochastic kernel $\pi(x, E)$ on $S \times \widehat{S}$ such that for each $\lambda > 0$, $x \in S$, and $E \in \mathscr{B}(\widehat{S})$,

$$(1.14) \qquad \boldsymbol{E}_x[e^{-\lambda\tau}, \boldsymbol{X}_\tau \in E \mid \boldsymbol{X}_{\tau-}] = \pi(\boldsymbol{X}_{\tau-}, E)\boldsymbol{E}_x[e^{-\lambda\tau} \mid \boldsymbol{X}_{\tau-}] \quad a.a. \text{ on } \{\tau < \infty\}.$$

Definition 1.3. Suppose a branching Markov process satisfies the Assumption 1.1. Then we shall call $\pi(x, E)$ the *branching law of* $\boldsymbol{X}$. Also, we shall say that *a branching process possesses a branching law* π if and only if it satisfies the Assumption 1.1 with the kernel π.

The existence of a branching law for a branching Markov process $\boldsymbol{X}$ is generally assured: for example, if $\boldsymbol{X}$ is a Hunt process with a reference measure, it possesses a branching law (cf. Theorem 0.2 of §0.2).

Remark 1.2. For a branching law, we can always assume $\pi(x, S) = 0$, $x \in S$. To give a branching law, it is equivalent to give the following system $(q_n(x), \pi_n(x, dy))_{n=0}^{\infty}$, (where $q_n(x) \in \boldsymbol{B}(S)^+$ and $\pi_n(x, dy)$ is a stochastic kernel on $S \times S^n$, $n = 0, 1, \cdots, \infty$) by the relation

$$(1.15) \qquad q_n(x) = \pi(x, S^n), \quad n = 0, 2, \cdots, +\infty,$$

$$(1.16) \qquad \pi_n(x, dy) = \pi(x, dy \cap S^n)/q_n(x), \quad n = 0, 2, \cdots, +\infty.^{(28)}$$

28) If $q_n(x) = 0$, take as $\pi_n(x, dy)$ any probability measure $\pi_n(dy)$ on S^n.

We shall call $(q_n(x), \pi_n(x, dy))_{n=0}^{\infty}$ the *branching system of the process* $\boldsymbol{X}$.

§1.2. Fundamental theorem.

We have formulated in the previous section the independence property of particles of a branching process in terms of its semigroup in the form (1.2). In this section, we shall give other mathematical formulations of the independence property and discuss their equivalence.

Let S, $\boldsymbol{S}=\bigcup_{n=0}^{\infty} S^n$, and $\widehat{\boldsymbol{S}}$ be as above. Let $\boldsymbol{X}=(\Omega, \mathcal{B}_t, \boldsymbol{X}_t, \boldsymbol{P}_{\boldsymbol{x}}, \theta_t)$

be a right continuous strong Markov process with Δ and ∂ as traps such that $\overline{\mathcal{B}}_{t+0}=\mathcal{B}_t$: and it is *not assumed apriori* to be a branching Markov process. Define ξ_t, τ, τ_n, $n=0, 1, 2, \cdots, +\infty$, in the same way as in the previous section.

Definition 1.4. (i) $\boldsymbol{X}$ is said to *satisfy the condition* (C. 1) if for every $x\in \boldsymbol{S}$,

(C. 1) $\quad \boldsymbol{P}_{\boldsymbol{x}}[\tau_\infty=e_\Delta,\ \tau_\infty<\infty]=\boldsymbol{P}_{\boldsymbol{x}}[\tau_\infty<\infty]$.

(ii) $\boldsymbol{X}$ is said to *satisfy the condition* (C. 2) if for every $x\in S$,

(C. 2) $\quad \boldsymbol{P}_{\boldsymbol{x}}[\tau=s]=0, \quad$ for all $s\geq 0$.

We note that if $\boldsymbol{X}$ is quasi-left continuous (cf. §0. 1. Definition 0. 3), then (C. 1) and (C. 2) are automatically satisfied. As for (C. 1), we have

$$\boldsymbol{P}_{\boldsymbol{x}}[\lim_{n\to\infty} \boldsymbol{X}_{\tau_n}=\boldsymbol{X}_{\tau_\infty},\ \tau_\infty<\infty]=\boldsymbol{P}_{\boldsymbol{x}}[\tau_\infty<\infty];$$

also, it is clear that if $\tau_\infty<\infty$ and $\lim_{n\to\infty} \boldsymbol{X}_{\tau_n}$ exists, then this limit must be Δ and consequently for $\boldsymbol{x}\in \boldsymbol{S}$ we have

$$\boldsymbol{P}_{\boldsymbol{x}}[\boldsymbol{X}_{\tau_\infty}=\Delta,\ \tau_\infty<\infty]=\boldsymbol{P}_{\boldsymbol{x}}[e_\Delta=\tau_\infty,\ \tau_\infty<\infty]=\boldsymbol{P}_{\boldsymbol{x}}[\tau_\infty<\infty].$$

As for (C. 2) we need only remark that $T\equiv s$ is an accessible Markov time.

Now we shall construct several stochastic processes from the process $\boldsymbol{X}=(\Omega, \mathcal{B}_t, \boldsymbol{X}_t, \boldsymbol{P}_{\boldsymbol{x}}, \theta_t)$.

(A) The process $\widetilde{X}=(\widetilde{\Omega}, \widetilde{\mathfrak{N}}_t, \widetilde{X}_t, \widetilde{P}_{\boldsymbol{x}}, \boldsymbol{x}\in \boldsymbol{S})$.

It is a stochastic process on $\widehat{\boldsymbol{S}}$ defined by (i)-(iv) below.

(i) Let $\Omega^{(n)}$ be the n-fold product of Ω and $\widetilde{\Omega}=\bigcup_{n=1}^{\infty} \Omega^{(n)}$ be their sum.

(ii) $\widetilde{X}_t(\tilde{\omega})$, $t\geqq 0$, $\tilde{\omega}\in\widetilde{\Omega}$, is defined for $\tilde{\omega}=(\omega_1, \omega_2, \cdots, \omega_n)\in\Omega^{(n)}$ by

$$\widetilde{X}_t(\tilde{\omega})=\gamma(\boldsymbol{X}_t(\omega_1), \boldsymbol{X}_t(\omega_2), \cdots, \boldsymbol{X}_t(\omega_n)), \tag{1.17}$$

where γ is defined in §0.2 (0.19).

(iii) $\widetilde{\mathfrak{N}}_t=\sigma(\widetilde{\Omega}, \mathcal{B}(\widehat{\boldsymbol{S}}); \widetilde{X}_s(\tilde{\omega}); s\leq t)$, $\widetilde{\mathfrak{N}}_\infty=\bigvee_{t>0}\widetilde{\mathfrak{N}}_t$.

(iv) For $A\in\widetilde{\mathfrak{N}}_\infty$,

$$\begin{aligned}(1.18)\quad \widetilde{P}_{\boldsymbol{x}}[A] &= \boldsymbol{P}_{x_1}\times\boldsymbol{P}_{x_2}\times\cdots\times\boldsymbol{P}_{x_n}[A\cap\Omega^{(n)}], && \text{if } \boldsymbol{x}=[x_1, \cdots, x_n]\in S^n\\ &= \boldsymbol{P}_{\boldsymbol{x}}[A\cap\Omega], && \text{if } \boldsymbol{x}=\partial \text{ or } \Delta.\text{[29]}\end{aligned}$$

$\widetilde{P}_{\boldsymbol{x}}$ is well defined by the following

Lemma 1.1. *For every* $A\in\widetilde{\mathfrak{N}}_\infty$, *the right-hand side of* (1.18) *is independent of the order of* $(x_1, x_2, \cdots, x_n)$ *and hence it depends only on* $\boldsymbol{x}=[x_1, x_2, \cdots, x_n]$.

Proof. It is sufficient to show that $\boldsymbol{E}_{x_1}\times\boldsymbol{E}_{x_2}\times\cdots\times\boldsymbol{E}_{x_n}[g_1(\widetilde{X}_{t_1}) g_2(\widetilde{X}_{t_2})\cdots g_q(\widetilde{X}_{t_q})]$ is invariant under the permutation of $(x_1, x_2, \cdots, x_n)$ for every $0\leq t_1<t_2<\cdots<t_q$ and $g_1, g_2, \cdots, g_q\in \boldsymbol{C}_0(\boldsymbol{S})$. For this, by virtue of Lemma 0.2, we can assume $g_i=\widehat{f}_i$, $f_i\in\boldsymbol{C}^*(S)$. Then, the assertion is clear since

$$\widehat{f}_1(\widetilde{X}_{t_1}(\tilde{\omega}))\widehat{f}_2(\widetilde{X}_{t_2}(\tilde{\omega}))\cdots\widehat{f}_q(\widetilde{X}_{t_q}(\tilde{\omega}))=\prod_{i=1}^{n}\prod_{j=1}^{q}\widehat{f}_j(\boldsymbol{X}_{t_j}(\omega_i)),$$

where $\tilde{\omega}=(\omega_1, \omega_2, \cdots, \omega_n)$, is invariant under the permutation of

29) From this definition it is clear that $\widetilde{X}$ has Δ and ∂ as traps.

$(\omega_1, \omega_2, \cdots, \omega_n)$.

(**B**) The process $X_n = (\Omega, X_t^{(n)}, \boldsymbol{P}_x, x \in S^n)$, $n = 1, 2, \cdots$: For each $n = 1, 2, \cdots$, X_n is a stochastic process which is obtained by stopping $\boldsymbol{X}$ at time τ on S^n. Therefore Ω, $\{\boldsymbol{P}_x\}$, $x \in S^n$, are the same as those of $\boldsymbol{X}$, and $X_t^{(n)}$ is defined by

(1.19) $$X_t^{(n)}(\omega) = \boldsymbol{X}_{t \wedge \tau(\omega)}(\omega), \quad t \geqq 0, \ \omega \in \Omega.$$

(**C**) The process $\widetilde{X}_n = (\Omega^{(n)}, \widetilde{X}_t^{(n)}, \widetilde{P}_x, x \in S^n)$, $n = 1, 2, \cdots$. For each $n = 1, 2, \cdots$, $\widetilde{X}_n$ is a stochastic process which is obtained from the $\widetilde{X}$ defined in (A) just as the process X_n is obtained from the process $\boldsymbol{X}$; i.e., $\widetilde{X}_n$ is the stopped process of $\widetilde{X}$ on S^n at time $\hat{\tau}$, where $\hat{\tau}$ is the first leaving time from S^n. To be precise $\Omega^{(n)}$ is the n-fold product of Ω, and $\widetilde{X}_t^{(n)}(\omega) = \widetilde{X}_{t \wedge \hat{\tau}(\tilde{\omega})}(\tilde{\omega}) = \gamma(X_{t \wedge \hat{\tau}(\tilde{\omega})}(\omega_1), \cdots, X_{t \wedge \hat{\tau}(\tilde{\omega})}(\omega_n))$ for $\tilde{\omega} = (\omega_1, \omega_2, \cdots, \omega_n)$, where $\hat{\tau}(\tilde{\omega}) = \min\{\tau(\omega_1), \tau(\omega_2), \cdots, \tau(\omega_n)\}$. $\{\widetilde{P}_x\}$ $x \in S^n$, is defined by (1.18).

Definition 1.5. (1) $\boldsymbol{X}$ is said to *have the property* B.I if it has the following

Property B.I. The processes $\boldsymbol{X}$ and $\widetilde{X}$ are equivalent.

(2) $\boldsymbol{X}$ is said to *have the property* B.II if it has the following

Property B.II. The processes X_n and $\widetilde{X}_n$ are equivalent for each n, $(n = 1, 2, \cdots)$.

(3) $\boldsymbol{X}$ is said to *have the property* B.III if it has the following

Property B.III. For $x \in S^n$, $n = 2, 3, \cdots$, we have

(i) $T_t^0 \hat{f}(x) = \widehat{(T_t^0 \hat{f})|_S}(x)$, for all $f \in \boldsymbol{B}^*(S)$, and

(ii) $\displaystyle\int_0^t \int_{\hat{S}} \hat{f}(s, y)\, \psi(x; ds\, dy) =$

$$= \int_0^t \langle T_s^0 f(s, \cdot) \mid (\int_S \hat{f}(s, y)\, \psi(\cdot\,; ds\, dy)) |_S \rangle \ (x)^{(30)}$$

---

30) $T_s^0 \hat{f}(s, \cdot)(x) = \boldsymbol{E}_x[\hat{f}(s, x_s);\ s < \tau]$, $x \in S$. The right-hand side is equal to. if $x = [x_1, x_2, \cdots, x_n]$, $\sum_{i=1}^n \int_0^t \left\{ \Pi_{j=1, j \neq i}^n T_s^0 \hat{f}(s, \cdot)(x_j) \right\} \int_S \hat{f}(s, y)\, \psi(x_i; ds\, dy)$, where $\hat{f}(s, y) = \prod_{j=1}^n f(s, y_j)$ if $y = [y_1, \cdots, y_n] \in S^n$.

for all $f \in \boldsymbol{B}^*([0, \infty) \times S)$, where

$$T_t^0 \widehat{f}(\boldsymbol{x}) = \boldsymbol{E}_{\boldsymbol{x}}[\widehat{f}(\boldsymbol{X}_t);\ t<\tau] \text{ and}$$
$$\psi(\boldsymbol{x};\ ds\, dy) = \boldsymbol{P}_{\boldsymbol{x}}[\boldsymbol{X}_\tau \in dy, \tau \in ds].$$

We have now the following fundamental

Theorem 1.2. *Let* $\boldsymbol{X} = (\Omega, \mathcal{B}_t, \boldsymbol{P}_{\boldsymbol{x}}, \boldsymbol{X}_t)$ *be a right continuous strong Markov process on* $\widehat{S}$ *such that* $\overline{\mathcal{B}}_{t+0} = \mathcal{B}_t$ *with* ∂ *and* Δ *as traps. Then it holds that*:
(a) $\boldsymbol{X}$ *is a branching Markov process if and only if it has the property* B.I;
(b) *If* $\boldsymbol{X}$ *has the property* B. I *then it has the property* B. II;
(c) *If* $\boldsymbol{X}$ *has the property* B. II *and satisfies the condition* (C. 2) *then it has the property* B. III;
(d) *If* $\boldsymbol{X}$ *has the property* B. III *and satisfies the condition* (C. 1) *and* (C. 2), *then it is a branching Markov process.*

Proof.

(i) Proof of (a): First we shall show that if $\boldsymbol{X}$ is a branching process, then it has the property B. I. For this, it is sufficient to show that for $0<t_1<t_2<\cdots<t_q$, $g_1, g_2, \cdots, g_q \in \boldsymbol{C}_0(\boldsymbol{S})$ we have

$$\boldsymbol{E}_{\boldsymbol{x}}[g_1(\boldsymbol{X}_{t_1}) g_2(\boldsymbol{X}_{t_2}) \cdots g_q(\boldsymbol{X}_{t_q})] = \widetilde{\boldsymbol{E}}_{\boldsymbol{x}}[g_1(\widetilde{X}_{t_1}) g_2(\widetilde{X}_{t_2}) \cdots g_q(\widetilde{X}_{t_q})].$$

By virtue of Lemma 0. 2 we can assume $g_i = \widehat{f_i}$ where $f_i \in \boldsymbol{C}^*(S)$. By the definition of $\widetilde{P}_{\boldsymbol{x}}$, when $\boldsymbol{x} \in [x_1, x_2, \cdots, x_n] \in S^n$,

$$\widetilde{E}_{\boldsymbol{x}}[\prod_{j=1}^q \widehat{f}_j(\widetilde{X}_{t_j})] = \prod_{i=1}^n \boldsymbol{E}_{x_i}[\prod_{j=1}^q \widehat{f}_j(\boldsymbol{X}_{t_j})],$$

and hence what we should prove is the following equality

$$(1.20) \qquad \boldsymbol{E}_{\boldsymbol{x}}[\prod_{j=1}^q \widehat{f}_j(\boldsymbol{X}_{t_j})] = \prod_{i=1}^n \boldsymbol{E}_{x_i}[\prod_{j=1}^q \widehat{f}_j(\boldsymbol{X}_{t_j})].$$

When $q=1$, (1. 20) is (1. 2) itself. Suppose (1. 20) is true for $q = 1, 2, \cdots, r-1$, then

$$\boldsymbol{E}_{\boldsymbol{x}}[\prod_{j=1}^r \widehat{f}_j(\boldsymbol{X}_{t_j})] = \boldsymbol{E}_{\boldsymbol{x}}[\prod_{j=1}^{r-1} \widehat{f}_j(\boldsymbol{X}_{t_j}) \boldsymbol{T}_{t_r - t_{r-1}} \widehat{f}_r(\boldsymbol{X}_{t_{r-1}})]$$

$$= \boldsymbol{E}_{\boldsymbol{x}}[\prod_{j=1}^{r-1} \widehat{f}_j(\boldsymbol{X}_{t_j}) \widehat{(\boldsymbol{T}_{t_r - t_{r-1}} \widehat{f}_r) |_S}(\boldsymbol{X}_{t_{r-1}})]$$

$$= \boldsymbol{E}_{\boldsymbol{x}}[\prod_{j=1}^{r-2} \widehat{f}_j(\boldsymbol{X}_{t_j}) \widehat{g}(\boldsymbol{X}_{t_{r-1}})]$$

$$= \prod_{i=1}^{n} \boldsymbol{E}_{x_i}[\prod_{j=1}^{r-2} \widehat{f}_j(\boldsymbol{X}_{t_j}) \widehat{g}(\boldsymbol{X}_{t_{r-1}})]$$

$$= \prod_{i=1}^{n} \boldsymbol{E}_{x_i}[\prod_{j=1}^{r} \widehat{f}_j(\boldsymbol{X}_{t_j})],$$

where we set $g = f_{r-1} \cdot (T_{t_r - t_{r-1}} \widehat{f}_r)|_S$. Therefore (1.20) holds for every q. Suppose, conversely, that $\boldsymbol{X}$ has the property B. I. Then, for $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]$,

$$\boldsymbol{E}_{\boldsymbol{x}}[\widehat{f}(\boldsymbol{X}_t)] = \widetilde{\boldsymbol{E}}_{\boldsymbol{x}}[\widehat{f}(\widetilde{\boldsymbol{X}}_t)] = \prod_{i=1}^{n} \boldsymbol{E}_{x_i}[\widehat{f}(\boldsymbol{X}_t)], \quad f \in \boldsymbol{C}^*(S),$$

which proves that $\boldsymbol{X}$ satisfies (1.2); i.e., $\boldsymbol{X}$ is a branching Markov process.

(ii) Proof of (b): Obvious.

(iii) Proof of (c): Assume $\boldsymbol{X}$ has the property B. II and satisfies the condition (C. 2). Then for $\boldsymbol{x} = [x_1, x_2, \cdots, x_n] \in S^n$ and $f \in \boldsymbol{B}^*(S)$,

$$\begin{aligned} &\boldsymbol{E}_{\boldsymbol{x}}[\widehat{f}(\boldsymbol{X}_t);\ t < \tau] \\ &= \boldsymbol{E}_{\boldsymbol{x}}[\widehat{f}(X_t^{(n)});\ t < \tau] \\ &= \widetilde{\boldsymbol{E}}_{\boldsymbol{x}}[\widehat{f}(\widetilde{X}_t^{(n)});\ t < \tilde{\tau}] \\ &= \boldsymbol{E}_{x_1} \times \cdots \times \boldsymbol{E}_{x_n}[\prod_{i=1}^{n} f(\boldsymbol{X}_t(\omega_i)) \cdot I_{\{\tau(\omega_i) > t\}}(\omega_i)] \\ &= \prod_{i=1}^{n} \boldsymbol{E}_{x_i}[f(\boldsymbol{X}_t);\ \tau > t]. \end{aligned}$$

Hence, the first condition of the property B. III is satisfied. Next, let $f \in \boldsymbol{B}^*([0, \infty) \times S)$. By the condition (C. 2) we have

$$\widetilde{P}_{\boldsymbol{x}}[\cup_{k,j=1,\ k \neq j}^{n} \{\widetilde{\omega} = (\omega_1, \omega_2, \cdots, \omega_n);\ \tau(\omega_k) = \tau(\omega_j)\}] = 0,$$

$\boldsymbol{x} \in S_n$, $n = 2, 3, \cdots$, and hence if $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]$,

$$\widehat{\boldsymbol{E}}_{\boldsymbol{x}}[\widehat{f}(\tau, \boldsymbol{X}_\tau);\ \tau \leq t]$$

$$=\boldsymbol{E}_x[\widehat{f}(\tau, X_\tau^{(n)});\ \tau\leq t]$$
$$=\widetilde{\boldsymbol{E}}_x[\widehat{f}(\hat{\tau}, \widetilde{X}_{\hat{\tau}}^{(n)});\ \hat{\tau}\leq t]$$
$$=\sum_{i=1}^n \boldsymbol{E}_{x_1}\times\cdots\times\boldsymbol{E}_{x_n}[\widehat{f}(\hat{\tau}, \widetilde{X}_{\hat{\tau}}^{(n)});\ \hat{\tau}(\tilde{\omega})=\tau(\omega_i)\leq t]$$
$$=\sum_{i=1}^n\int\cdots\int\boldsymbol{P}_{x_1}\times\cdots\times\boldsymbol{P}_{x_n}[d\omega_1\cdots d\omega_n]\,[\Pi_{\substack{j=1\\ j\neq i}}^n\{\widehat{f}(\tau(\omega_i), X_{\tau(\omega_i)}(\omega_j))$$
$$\cdot I_{\{\tau(\omega_i)<\tau(\omega_j)\}}\}\cdot\widehat{f}(\tau(\omega_i), \boldsymbol{X}_{\tau(\omega_i)}(\omega_i))\cdot I_{\{\tau(\omega_i)\leq t\}}]$$
$$=\sum_{i=1}^n\int_{\boldsymbol{S}}\int_0^t\{\Pi_{\substack{j=1\\ j\neq i}}^n\boldsymbol{E}_{x_j}[\widehat{f}(s, \boldsymbol{X}_s);\ s<\tau]\cdot\widehat{f}(s, \boldsymbol{y})\}$$
$$\cdot\boldsymbol{P}_{x_i}[\boldsymbol{X}_\tau\in d\boldsymbol{y},\ \tau\in ds],$$

which proves that $\boldsymbol{X}$ satisfies the second condition of the property B. III.

(iv) Proof of (d): We shall prove a proposition which includes Theorem 1.2, (d) as a special case.

Let $T_t^0(x, dy)$ be a kernel on $([0, \infty)\times S)\times S$ satisfying

(i) $T_t^0(x, \cdot)$ is a signed Borel measure on S with finite total variation for every $(t, x)\in[0, \infty)\times S$,

(ii) $T_t^0(\cdot, B)$ is Borel measurable on $[0, \infty)\times S$ for every $B\in\mathscr{B}(S)$ and

(iii) $T_{t+s}^0(x, dy)=\int_S T_s^0(x, dz)T_t^0(z, dy)$ for every $t, s\in[0, \infty)$, $x\in S$. Set $T_t^0f(x)=\int_S T_t^0(x, dy)f(y)$.

Let $\psi(x;\ ds\,d\boldsymbol{y})$ be another kernel on $S\times\{[0, \infty)\times(\boldsymbol{S}-S)\}$ satisfying

(i) $\psi(x;\ \cdot)$ is a signed measure on $[0, \infty)\times(\boldsymbol{S}-S)$ with finite total variation on $[0, t]\times(\boldsymbol{S}-S)$ for every $t>0$,

(ii) $\psi(\cdot\,;\ \Gamma)$ is measurable on S for any Borel subset Γ of $[0, \infty)\times(\boldsymbol{S}-S)$ and

(iii) $|\psi|(y;\ \{s\}\times(\boldsymbol{S}-S))=0$ for every $y\in S$ and $s>0$, where $|\psi|$ denotes the total variation of ψ.

Furthermore, we assume that T_t and ψ satisfy

$$(1.21)\quad \int_S T_t^0(x, dy)\int_0^s\int_{\boldsymbol{S}-S}\psi(y;\ dr\ d\boldsymbol{z})f(\boldsymbol{z})=\int_t^{t+s}\int_{\boldsymbol{S}-S}\psi(x;\ dr\ d\boldsymbol{z})f(\boldsymbol{z}),$$

for every $f \in \boldsymbol{B}(S-S)$.

Now we extend T_t and ψ to kernels $\boldsymbol{T}_t(x, dy)$ and $\psi(x; dt\, dy)$, $x, y \in \boldsymbol{S}$, $t \in [0, \infty)$ by

$$(1.22) \qquad \int_{\boldsymbol{S}} \boldsymbol{T}_t^{(0)}(x, dy)\widehat{f}(y) = \widehat{T_t^0 f}(x), \quad f \in \overline{\boldsymbol{B}}^*(S)^+$$

$$(1.23) \qquad \int_0^t \int_{\boldsymbol{S}} \psi(x; ds\, dy)\widehat{f}(s, y) = \int_0^t \langle T_s^0 f(s, \cdot) \Big| \int_{S-S} \psi(\cdot\,; ds\, dz)\widehat{f}(s, z)\rangle \;(x)$$

(cf. Lemma 0.3). The support of $\boldsymbol{T}_t^{(0)}(x, dy)$ is concentrated on S^n when $x \in S^n$ and it defines a semi-group on $\boldsymbol{B}(\boldsymbol{S})$.

We define kernels $\psi^{(n)}(x; dt\, dy)$ $(n=0, 1, 2, \cdots)$ on $\boldsymbol{S} \times ([0, \infty) \times \boldsymbol{S})$ by

$$\Phi^{(0)}(x; t, dy) = \delta_{\{x\}}\,(dy)$$

$$(1.24) \qquad \Phi^{(1)}(x; t, dy) = \int_0^t \psi(x; ds\, dy)$$

$$\Phi^{(k)}(x; t, dy) = \int_0^t \int_{\boldsymbol{S}} \psi(x; dv\, dz)\Phi^{(k-1)}(z; t-v, dy)$$

and

$$\psi^{(k)}(x; dt\, dy) = d_t \Phi^{(k)}(x; t, dy).$$

Moreover set for each $k=0, 1, 2, \cdots$

$$(1.25) \qquad \boldsymbol{T}_t^{(k)}(x, dy) = \int_0^t \int_{\boldsymbol{S}} \psi^{(k)}(x; ds\, dz)\, \boldsymbol{T}_{t-s}^{(0)}(z, dy)$$

and

$$(1.26) \qquad \boldsymbol{T}_t^{(k)} f(x) = \int_{\boldsymbol{S}} \boldsymbol{T}_t^{(k)}(x, dy) f(y), \quad f \in \boldsymbol{B}(\boldsymbol{S}).$$

Proposition 1.3. *When* $\sum_{k=0}^{\infty} \boldsymbol{T}_t^{(k)} \widehat{f}(x)$, $f \in \boldsymbol{B}(S)$, *converges, set*

$$(1.27) \qquad \boldsymbol{T}_t \widehat{f}(x) = \sum_{k=0}^{\infty} T_t^{(k)} \widehat{f}(x).$$

Then $\boldsymbol{T}_t \widehat{f}(x)$ *has the branching property*:

$$(1.28) \qquad \boldsymbol{T}_t \widehat{f}(x) = \widehat{(\boldsymbol{T}_t \widehat{f}\,|_S)}(x), \quad x \in \boldsymbol{S}.$$

Proof of the proposition consists of several steps.

Lemma 1.2. *$T_t^{(n)}$ and $\psi^{(n)}$ satisfy the following relations for $f\in B(S)$ and $0\leq k\leq n$;*

$$\int_t^{t+s}\int_S \psi(x;\, dr\, dy)f(y)=T_t^{(0)}\left[\int_0^s\int_S \psi(\cdot\,;\, dr\, dy)f(y)\right](x) \tag{1.29}$$

$$\Phi^{(n)}(t)f(x)=\int_0^t \psi^{(n-k)}(dr)\Phi^{(k)}(t-r)f(x)^{(31)} \tag{1.30}$$

$$T_t^{(n)}f(x)=\int_0^t \psi^{(n-k)}(dr)\,T_{t-r}^{(k)}f(x) \tag{1.31}$$

$$T_v^{(0)}T_{t-v}^{(n)}f(x)=\int_v^t \psi(dr)\,T_{t-r}^{(n-1)}f(x). \tag{1.32}$$

Proof. By (1.23) we have

$$\int_t^{t+s}\int_S \psi(x;\, dr\, dy)\hat{g}(y)=\int_t^{t+s}\left\langle T_r^0 g\,\middle|\,\int_S \psi(\cdot\,;\, dr\, dy)\hat{g}(y)\right\rangle(x)$$

and by (1.21) this is equal to

$$\int_0^s\left\langle T_{r+t}^0 g\,\middle|\,\int_S \psi(\cdot\,;\, dr+t, dy)\hat{g}(y)\right\rangle(x)$$

$$=\int_0^s\left\langle T_t^0 T_r^0 g\,\middle|\,\int_S T_t^0\psi(\cdot\,;\, dr\, dy)\hat{g}(y)\right\rangle(x)$$

$$=T_t^0\int_0^s\left\langle T_r^0 g\,\middle|\,\int_S \psi(\cdot\,;\, dr\, dy)\hat{g}(y)\right\rangle(x)$$

$$=T_t^0\left[\int_0^s\int_S \psi(\cdot\,;\, dr\, dy)\hat{g}(y)\right](x)^{(32)}.$$

Thus (1.29) is proved if f is of the form $\hat{g}$, $g\in C^*(S)$. By virtue of Lemma 0.2, (1.29) holds for every $f\in B(S)$.

---

31) $\Phi^{(n)}(t)f(x)\equiv\int_S \Phi^{(n)}(x;\, t,\, dy)f(y)$,

$\psi^{(n-k)}(dr)f(x)=\int_S \psi^{(n-k)}(x;\, dr\, dy)f(y)$. $f\in B(S)$.

32) $T_t^0\langle f\,|\,g\rangle=\langle T_t^0 f\,|\,T_t^0 g\rangle$.
Indeed, by (0.36)

$$T_t^0\langle f\,|\,g\rangle=\lim_{\epsilon\downarrow 0}T_t^0\{(\widehat{f+\epsilon g}-\hat{f})/\epsilon\}=\lim_{\epsilon\downarrow 0}[\widehat{T_t^0(f+\epsilon g)}-\widehat{T_t^0 f}]/\epsilon$$

$$=\lim_{\epsilon\downarrow 0}(\widehat{T_t^0 f+\epsilon T_t^0 g}-\widehat{T_t^0 f})/\epsilon=\langle T_t^0 f\,|\,T_t^0 g\rangle.$$

(1. 30) is a usual formula for iteration of convolutions. (1. 31) follows from (1. 25) and (1. 30).

Now

$$\boldsymbol{T}_v^{(0)}\boldsymbol{T}_{t-v}^{(0)}f(\boldsymbol{x})=\boldsymbol{T}_v^{(0)}\left\{\int_0^{t-v}\psi(dr)\,\boldsymbol{T}_{t-v-r}^{(n-1)}f\right\}(\boldsymbol{x})$$
$$=\int_0^{t-v}\boldsymbol{T}_v^0\psi(dr)\,\boldsymbol{T}_{t-v-r}^{(n-1)}f(\boldsymbol{x})$$

and by (1. 29) this equals

$$\int_0^{t-v}d_r\Phi(r+v)\,\boldsymbol{T}_{t-r-v}^{(n-1)}f(\boldsymbol{x})=\int_v^t\psi(dr)\,\boldsymbol{T}_{t-r}^{(n-1)}f(\boldsymbol{x}).$$

which proves (1. 31).

Lemma 1.3. *For* $m\geq n-1$, $m\neq n$, $\boldsymbol{x}=[x_1, x_2, \cdots, x_n]\in S^n$, *we have*

$$\text{(1. 33)}\qquad \int_0^t\int_{S^m}\psi(\boldsymbol{x};\, ds\,d\boldsymbol{y})\widehat{f}(s, \boldsymbol{y})$$
$$=\int_0^t\langle T_s^0 f(s,\ \cdot)\,|\int_{S^{m-n+1}}\psi(\cdot\,;\, ds\,d\boldsymbol{y})\widehat{f}(s, \boldsymbol{y})\rangle(\boldsymbol{x}),$$

for every $f\in\boldsymbol{B}^*([0, \infty)\times S)$.

Proof. For $0<\lambda<1$, we have from (1. 23),

$$\int_0^t\int_S\psi(\boldsymbol{x};\, ds\,d\boldsymbol{y})\widehat{\lambda f}(s, \boldsymbol{y})=\sum_{m=0}^{\infty}\lambda^m\int_0^t\int_{S^m}\psi(\boldsymbol{x};\, ds\,d\boldsymbol{y})\widehat{f}(s, \boldsymbol{y})$$
$$=\int_0^t\langle T_r^0(\lambda f(s,\ \cdot))\,|\int_S\psi(\cdot\,;\, ds\,d\boldsymbol{y})\widehat{\lambda f}(s, \boldsymbol{y})\rangle\ (\boldsymbol{x})$$
$$=\sum_{r=0}^{\infty}\lambda^{n-1}\lambda^r\int_0^t\langle T_s^0 f(s,\ \cdot)\,|\int_{S^r}\psi\ (\cdot\,;\, ds\,d\boldsymbol{y})\widehat{f}(s, \boldsymbol{y})\rangle\ (\boldsymbol{x}).$$

Comparing the coefficients of λ^m we have (1. 33).

Remark 1.3. By (1. 33) we see that, if $\boldsymbol{x}\in S^n$, $\psi(\boldsymbol{x};\, ds\,d\boldsymbol{y})$ has no mass on S^k for $k<n-1$ and $k=n$ with respect to $d\boldsymbol{y}$.

Lemma 1.4. *For every* $\boldsymbol{x}=[x_1, x_2, \cdots, x_n]\in S^n$, $m\geq n-1$, $m\neq n$, *and* $f_1(s, x), \cdots, f_m(s, x)\in\boldsymbol{B}([0, \infty)\times S)$, *we have*

$$(1.34)\qquad \int_0^t\int_{S^m}\psi(x;\,ds\,dy)\left\{\frac{1}{m!}\sum_{\pi}\prod_{j=1}^m f_j(s,y_{\pi(j)})\right\}$$
$$=\int_0^t\int_{S^m}\psi(x;\,ds\,dy)\left\{\frac{1}{m!}\sum_{\pi}\prod_{j=1}^m f_{\pi(j)}(s,y_j)\right\}$$
$$=\int_0^t\sum_{i=1}^n\frac{1}{{}_mC_{n-1}}\sum_{(q_1,q_2,\cdots,q_{m-n+1})}\frac{1}{(n-1)!}\sum_{\hat{\pi}}^{(\hat{q})}\int_{S^{m-n+1}}\psi(x_i;\,ds\,dy^0)$$
$$\times\left\{\frac{1}{(m-n+1)!}\sum_{\pi}^{(q)}\left[\prod_{h=1}^{m-n+1}f_{q_{\pi(h)}}(s,y_h^0)\right]\right\}$$
$$\times\prod_{j\neq i}T_s^0 f_{\hat{q}_{\hat{\pi}(j)}}(s,\cdot)(x_j),^{(33)}$$

where $\sum_{\pi}$ *denotes the sum over all permutations* π *on* $(1,2,,\cdots,m)$, $\sum_{(q_1,\cdots,m-n+1)}$ *the sum over all ordered choices* $(q_1,q_2,\cdots,q_{m-n+1})$ *from* $(1,2,\cdots,m)$, $\sum_{\pi}^{(q)}$ *the sum over all permutations* π *on* $(q_1,\cdots,q_{m-n+1})$, *and* $\sum_{\hat{\pi}}^{(\hat{q})}$ *the sum over all permutations* $\hat{\pi}$ *on* $(\hat{q}_1,\cdots,\hat{q}_{n-1})$ *which is the remainder of* $(1,2,\cdots,m)$ *excluding* $(q_1,\cdots,q_{m-n+1})$.

Proof. Let $f_i\in \boldsymbol{B}([0,\infty)\times S)$, $i=1,2,\cdots,m$, then by a combinatorial lemma on permanents,[34]

$$\int_0^t\int_{S^m}\psi(x;\,ds\,dy)\left\{\sum_{\pi}\prod_{j=1}^m f_{\pi(j)}(s,y_j)\right\}$$
$$=\int_0^t\int_{S^m}\psi(x;\,ds\,dy)\left\{\widehat{\left(\sum_{q=1}^m f_q\right)}(s,y)\right.$$
$$\left.-\sum_{(k_1,\cdots,k_{m-1})}\widehat{\left(\sum_{q=1}^{m-1}f_{k_q}\right)}(s,y)+\cdots+(-1)^{m-1}\sum_{k=1}^m\hat{f}_k(s,y)\right\},$$

and by (1.33) this is equal to

$$\int_0^t\sum_{i=1}^n\int_{S^{m-n+1}}\psi(x_i;\,ds\,dy)\left\{\widehat{\left(\sum_{q=1}^m f_q\right)}\right\}(s,y)\cdot\prod_{j\neq i}T_s^0\left(\sum_{q=1}^m f_q(s,\cdot)\right)(x_j)$$
$$-\sum_{(k_1,\cdots,k_{m-1})}\widehat{\left(\sum_{q=1}^{m-1}f_{k_q}\right)}(s,y)\cdot\prod_{j\neq i}T_s^0\left(\sum_{q=1}^{m-1}f_{k_q}(s,\cdot)\right)(x_j)$$
$$\left.+\cdots\right\}$$

Now applying again the lemma on permanents to the integrand { },

33) $y=[y_1,y_2,\cdots,y_n]$, $y^0=[y_1^0,y_2^0,\cdots,y_{m-n+1}^0]$.
34) Cf. Ryser [38].

the above expression becomes,

$$\int_0^t \sum_{i=1}^n \int_{S^{m-n+1}} \psi(x_i;\, ds\, dy) \Big\{ \sum_\pi \prod_{r=1}^{m-n+1} f_{\pi(r)}(s, y_r) \cdot \prod_{j\neq i} T_s^0 f_{\pi(r_j)}(s, \cdot)\ (x_j) \Big\},$$

where $r=1, 2, \cdots, m-n+1$, $\{r_j;\ 1\leq \underset{j\neq i}{j} \leq n\} = \{m-n+2, m-n+3, \cdots, m\}$, and π is a permutation on $\{1, 2, \cdots, m\}$. Since

$$\sum_\pi = \sum_{(q_1, \cdots, q_{m-n+1})} \sum_{\hat{\pi}}^{(\hat{q})} \sum_\pi^{(q)},$$

this is equal to

$$\int_0^t \sum_{i=1}^n \sum_{(q_1, \cdots, q_{m-n+1})} \sum_{\hat{\pi}}^{(\hat{q})} \sum_\pi^{(q)} \int_{S^{m-n+1}} \psi(x_i;\, ds\, dy) \cdot \sum_{r=1}^{m-n+1} f_{q_{\pi(r)}}(s, y_r) \cdot \prod_{j\neq i} T_s^0 f_{\hat{q}_{\hat{\pi}(j)}}(s, \cdot)(x_j).$$

Now (1.34) is clear since $m! = {}_mC_{n-1}(n-1)!(m-n+1)!$.

Lemma 1.5. *For* $x = [x_1, x_2, \cdots, x_n] \in S^n$,

$$(1.35)\qquad \sum_{r_1+\cdots+r_n=r} \int_0^t \sum_{i=1}^n \int_S \psi(x_i;\, ds\, dy)\, T_{t-s}^{(r_i)} \hat{f}(y) \prod_{j\neq i} T_s^{(0)} T_{t-s}^{(r_j)} \hat{f}(x_j) = \sum_{r_1+\cdots+r_n=r+1} \prod_{j=1}^n T_t^{(r_j)} \hat{f}(x_j)^{(35)},$$

for every $f \in \boldsymbol{B}^*(S)$.

Proof. Set $T_v^{(0)} T_{t-v}^{(r_j)} \hat{f}(x_j) = g^{(r_j)}(v)$; then if $r_j=0$, $g^{(r_j)}$ is independent of v by the semi-group property of $T_s^{(0)}$. If $r_j \geq 1$, then by (1.32)

$$g^{(r_j)}(v) = \int_v^t \int_S \psi(x_j;\, ds\, dy)\, T_{t-s}^{(r_j-1)} \hat{f}(y).$$

Hence the left-hand side of (1.35) is equal to

$$\sum_{r_1+\cdots+r_n=r} \int_0^t \sum_{k=1}^n d_v(-g^{(r_k+1)}(v)) \prod_{j\neq k} g^{(r_j)}(v).$$

35) $\sum_{r_1+r_2+\cdots+r_n=k}$ denotes the sum over all $(r_1, r_2, \cdots, r_n)$ such that $r_i \geq 0$, $i=1, 2, \cdots, n$ and $r_1+r_2+\cdots r_n=k$.

Writing r_k+1 as r_k and noting $dg^{(0)}(v)\equiv 0$, this equals

$$\sum_{r_1+\cdots+r_n=r+1}\int_0^t\sum_{k=1}^n d_v(-g^{(r_k)}(v))\cdot\prod_{j\neq k}g^{(r_j)}(v)$$
$$=\sum_{r_1+\cdots+r_n=r+1}\prod_{j=1}^n g^{(r_j)}(0)=\sum_{r_1+\cdots+r_n=r+1}\prod_{j=1}^n T_t^{(r_j)}\widehat{f}(x_j).$$

Lemma 1.6. *For* $x=[x_1, x_2, \cdots, x_n]\in S^n$ *and* $r\geq 0$ *we have*

$$(1.36)\qquad T_t^{(r)}\widehat{f}(x)=\sum_{r_1+r_2+\cdots+r_n=r}\prod_{j=1}^n T_t^{(r_j)}\widehat{f}(x_j),$$

for every $f\in B^*(S)$.

Proof. First we note the following identity which can be easily verified. If $F=F(n_1, n_2, \cdots, n_k)$, ($n_i$; a non-negative integer),

$$(1.37)\qquad \sum_{n_1+n_2+\cdots+n_k=n}F(n_1, n_2, \cdots, n_k)$$
$$=\sum_{n_1+n_2+\cdots+n_k=n}\frac{1}{k!}\sum_{\pi}F(n_{\pi(1)}, n_{\pi(2)}, \cdots, n_{\pi(k)}).$$

We shall prove (1.36) by induction on r. If $r=0$, then (1.36) is just (1.22). Now assume (1.36) is true for $r=1, 2, \cdots, r$, then using (1.31), the induction hypothesis, and (1.37) successively, we have

$$T_t^{(r+1)}\widehat{f}(x_j)=\int_0^t\int_S\psi(x; ds\, dy)\,T_{t-s}^{(r)}\widehat{f}(y)$$
$$=\sum_{m=0}^\infty\int_0^t\int_{S^m}\psi(x; ds\, dy)\Big[\sum_{r_1+\cdots+r_m=r}\prod_{j=1}^m T_{t-s}^{(r_j)}\widehat{f}(y_j)\Big]^{(36)}$$
$$=\sum_{m=0}^\infty\int_0^t\int_{S^m}\psi(x; ds\, dy)\Big[\sum_{r_1+\cdots+r_m=r}\frac{1}{m!}\sum_\pi\prod_{j=1}^m T_{t-s}^{(r_j)}\widehat{f}(y_{\pi(j)})\Big]$$
$$=\sum_{m=0}^\infty\sum_{r_1+\cdots+r_m=r}\int_0^t\int_{S^m}\psi(x; ds, dy)\Big[\frac{1}{m!}\sum_\pi\prod_{j=1}^m T_{t-s}^{(r_j)}\widehat{f}(y_{\pi(j)})\Big].$$

Put $f_i(s, y)\equiv T_{t-s}^{(r_i)}\widehat{f}(y)$; then by (1.34) the above equals

$$\sum_{m=0}^\infty\sum_{r_1+\cdots+r_m=r}\int_0^t\sum_{i=1}^n\frac{1}{{}_mC_{n-1}}\sum_{(q_1,\cdots,q_{m-n+1})}\frac{1}{(n-1)!}\sum_\pi^{(\hat{q})}$$
$$\int_{S^{m-n+1}}\psi(x_i; ds\, dy^0)\cdot\Big\{\frac{1}{(m-n+1)!}\sum_\pi^{(q)}\prod_{h=1}^{m-n+1}f_{q_{\pi(h)}}(s, y_h^0)\Big\}$$

36) $y=[y_1, y_2, \cdots, y_m]$.

$$\cdot \prod_{j \neq i} T_s^0 f_{\hat{q}\hat{\pi}(j)}(s, \cdot)(x_j)^{(37)}$$

$$= \sum_{m=0}^{\infty} \sum_{i=1}^{n} \int_0^t \frac{1}{{}_mC_{n-1}} \sum_{(q_1, \cdots, q_{m-k+1})} \Sigma^{(1)} \frac{1}{(n-1)!} \sum_{\hat{\pi}}^{(\hat{q})}$$

$$\int_{S^{m-n+1}} \psi(x_i; ds dy^0) \cdot \left\{ \Sigma^{(2)} \frac{1}{(m-n+1)!} \sum_{\pi}^{(q)} \prod_{h=1}^{m-n+1} f_{q\pi(h)}(s, y_h^0) \right\}$$

$$\cdot \prod_{j \neq i} T_s^0 f_{\hat{q}\hat{\pi}(j)}(s, \cdot)(x_j).^{(38)}$$

By applying (1. 37) and the induction hypothesis in { }, the above line becomes

$$\sum_{m=0}^{\infty} \sum_{i=1}^{n} \int_0^t \frac{1}{{}_mC_{n-1}} \sum_{(q_1, \cdots, q_{m-n+1})} \Sigma^{(1)} \frac{1}{(n-1)!} \sum_{\hat{\pi}}^{(\hat{q})}$$

$$\int_{S^{m-n+1}} \psi(x_i; ds\, dy^0) \cdot \left\{ \Sigma^{(2)} \prod_{h=1}^{m-n+1} f_{q_h}(s, y_h^0) \right\} \prod_{j \neq i} T_s^0 f_{\hat{q}\hat{\pi}(j)}(s, \cdot)(x_j)$$

$$= \sum_{m=0}^{\infty} \sum_{i=1}^{n} \int_0^t \frac{1}{{}_mC_{n-1}} \sum_{(q_1, \cdots, q_{m-n+1})} \Sigma^{(1)} \frac{1}{(n-1)!} \sum_{\hat{\pi}}^{(\hat{q})}$$

$$\int_{S^{m-n+1}} \psi(x_i; ds\, dy^0) \cdot T_{t-s}^{(r^*)} \hat{f}(y^0) \prod_{j \neq i} T_s^0 f_{\hat{q}\hat{\pi}(j)}(s, \cdot)(x_j)$$

$$= \sum_{m=0}^{\infty} \sum_{i=1}^{n} \int_0^t \sum_{r_1 + \cdots + r_n = r} \int_{S^{m-n+1}} \int \psi(x_i; ds\, dy^0)\, T_{t-s}^{(r_i)} \hat{f}(y^0)$$

$$\cdot \prod_{j \neq i} T_s^0 (T_{t-s}^{(r_j)} f)(x_j) \qquad \text{(by (2. 32))}$$

$$= \sum_{r_1 + \cdots + r_n = r+1} \prod_{j=1}^{n} T_t^{(r_j)} \hat{f}(x_j) \qquad \text{(by (2. 30))}.$$

Thus (1. 36) is proved for every r.

Now we are ready to complete the

Proof of Proposition 1. 3. By (1. 36), when $x = [x_1, x_2, \cdots, x_n] \in S^n$,

$$\boldsymbol{T}_t \hat{f}(\boldsymbol{x}) = \sum_{r=0}^{\infty} \boldsymbol{T}_t^{(r)} \hat{f}(\boldsymbol{x})$$

---

37) $y^0 = [y_1^0, y_2^0, \cdots, y_{m-n+1}^0]$.

38) $\Sigma^{(1)}$ is (for a fixed $(q_1, \cdots, q_{m-n+1})$) the sum over all $(r_{\hat{q}_1}, \cdots, r_{\hat{q}_{n-1}}, r^*)$ such that $r_{\hat{q}_1} + r_{\hat{q}_2} + \cdots + r_{\hat{q}_{n-1}} + r^* = r$ and $\Sigma^{(2)}$ is (for fixed $(q_1, \cdots, q_{m-n+1})$ and r^* such that $0 < r^* \leq r$) the sum over all $(r_{q_1}, r_{q_2}, \cdots r_{q_{m-n+1}})$ such that $r_{q_1} + r_{q_2} + \cdots + r_{q_{m-n-1}} = r^*$, where r^* and r_i are all nonnegative integers.

$$=\sum_{r=0}^{\infty}\sum_{r_1+\cdots+r_n=r}\prod_{j=1}^{n}\boldsymbol{T}_t^{(r_j)}\widehat{f}(x_j)$$

$$=\prod_{j=1}^{n}(\sum_{r=0}^{\infty}\boldsymbol{T}_t^{(r)}\widehat{f}(x_j))$$

$$=\prod_{j=1}^{n}\boldsymbol{T}_t\widehat{f}(x_j),$$

proving the branching property.

In order to complete the proof of (d), we set for a given strong Markov process $\boldsymbol{X}$ on $\widehat{S}$ with the Property B. III, (C. 1) and (C. 2),

$$T_t^0 f(x)=\boldsymbol{E}_x[f(\boldsymbol{X}_t);\ t<\tau],$$

$$\psi(x;\ dr\ dy)=\boldsymbol{P}_x[\tau\in dr,\ \boldsymbol{X}_\tau\in dy].$$

Then, as is easily seen, T_t^0 and Ψ satisfy all the conditions (i), (ii) and (iii) above. Note that (1. 21) is a simple consequence of the Markov property. It is easy to see by Property B. III and the strong Markov property that $\boldsymbol{T}_t^{(k)}$ and $\psi^{(k)}$ defined above are exactly

$$\boldsymbol{T}_t^{(k)}f(\boldsymbol{x})=\boldsymbol{E}_{\boldsymbol{x}}[f(\boldsymbol{X}_t);\ \tau_k\leq t<\tau_{k+1}],$$

$$\psi^{(k)}(\boldsymbol{x};\ dr\ d\boldsymbol{y})=\boldsymbol{P}_{\boldsymbol{x}}[\tau_k\in dr,\ \boldsymbol{X}_{\tau_k}\in d\boldsymbol{y}],\ k=0,1,2,\cdots.$$

By (C. 1),

$$\boldsymbol{T}_t\widehat{f}(\boldsymbol{x})=\boldsymbol{E}_{\boldsymbol{x}}[\widehat{f}(\boldsymbol{X}_t)]$$

$$=\sum_{k=1}^{\infty}\boldsymbol{E}_{\boldsymbol{x}}[\widehat{f}(\boldsymbol{X}_t);\ \tau_k\leq t<\tau_{k+1}]$$

$$=\sum_{k=0}^{\infty}\boldsymbol{T}_t^{(k)}\widehat{f}(\boldsymbol{x})$$

and hence $\boldsymbol{T}_t$ has the branching property by Proposition 1. 3.

Corollary. *Let* $\boldsymbol{X}$ *and* $\boldsymbol{X}'$ *be two branching Markov processes which satisfy the condition* (C. 1) *and* (C. 2) *and possess the branching law. If their non-branching parts and branching laws coincide,*[39] *then they are equivalent.*

39) To be precise, the non-branching parts are equivalent and $E_x^0[\pi(X_{\zeta^0-}^0,\ E);\ \zeta^0\leq t]$ defines the same measure on $\widehat{S}-S$ for every $E\in\mathcal{B}(\widehat{S}-S)$, $x\in S$, and $t\geq 0$, where (X_t^0, P_x^0, ζ^0) denotes the non-branching part. (In this case we shall say also that π and π' are equivalent.)

Proof. We have $T_t^0 f(x) = T_t'^0 f(x)$ for any $f \in \boldsymbol{B}(S)$ and $x \in S$. Since $\boldsymbol{X}$ and $\boldsymbol{X}'$ have the property B. III, we have from the property B. III (i),

$$T_t^0 f(\boldsymbol{x}) = T_t'^0 f(\boldsymbol{x}) \qquad \text{for every } f \in \boldsymbol{B}(\boldsymbol{S}) \text{ and } \boldsymbol{x} \in \boldsymbol{S}.$$

Now for $x \in S$

$$(1.38) \qquad \begin{aligned} \Phi^{(1)}(x;\, t, d\boldsymbol{y}) &= \boldsymbol{P}_x[\tau \leq t, \boldsymbol{X}_\tau \in d\boldsymbol{y}] \\ &= \boldsymbol{E}_x[\pi(\boldsymbol{X}_{\tau-}, d\boldsymbol{y}) \cdot \boldsymbol{P}_x[\tau \leq t \mid \boldsymbol{X}_{\tau-}]] \\ &= \boldsymbol{E}_x[\pi(\boldsymbol{X}_{\tau-}, d\boldsymbol{y});\, \tau \leq t] \\ &= E_x^0[\pi(X_{\zeta^0-}^0, d\boldsymbol{y});\, \zeta^0 \leq t]. \end{aligned}$$

Since the non-branching parts X^0 and X'^0 are equivalent and so are π and π', we have $\Phi^{(1)}(x;\, t, d\boldsymbol{y}) \equiv \Phi'^{(1)}(x;\, t, d\boldsymbol{y})$. Then by the property B. III (ii), $\psi(\boldsymbol{x};\, ds\, d\boldsymbol{y}) \equiv \psi'(\boldsymbol{x};\, ds\, d\boldsymbol{y})$. Therefore, the following relation

$$T_t^{(n)} f(\boldsymbol{x}) = \int_0^t \int_{\boldsymbol{S}} \psi(\boldsymbol{x};\, ds\, d\boldsymbol{y})\, T_{t-s}^{(n-1)} f(\boldsymbol{y}), \quad f \in \boldsymbol{B}(\boldsymbol{S}),$$

implies $T_t^{(n)} \equiv T_t'^{(n)}, n = 0, 1, 2, \cdots$ by induction, and hence

$$\boldsymbol{T}_t f(\boldsymbol{x}) = \sum_{n=0}^{\infty} \boldsymbol{T}_t^{(n)} f(\boldsymbol{x}) = \sum_{n=0}^{\infty} \boldsymbol{T}_t'^{(n)} f(\boldsymbol{x}) = \boldsymbol{T}_t' f(\boldsymbol{x}),$$

which proves the assertion of the corollary.

By this corollary, if $\boldsymbol{X}$ is a branching Markov process on $\hat{\boldsymbol{S}}$ which satisfies the condition (C. 1) and (C. 2) and possesses the branching law π, then $\boldsymbol{X}$ is uniquely determined by the non-branching part X^0 and the branching law π.

Definition 1. 6. Let $\boldsymbol{X}$[40] be a branching Markov process which satisfies the conditions (C. 1) and (C. 2) and possesses the branching law π. Let X^0 be the non-branching part of $\boldsymbol{X}$. Then we shall call $\boldsymbol{X}$ the *(X^0, π)-branching Markov process.*

40) It should be remembered that we are always assuming that $\boldsymbol{X} = (\boldsymbol{X}_t, \mathscr{B}_t)$ is strong Markov such that $\overline{\mathscr{B}}_{t+0} = \mathscr{B}_t$.

Under certain general conditions every branching Markov process is given as an (X^0, π)-branching Markov process. In fact, we have noticed already that (C. 1) and (C. 2) are satisfied if $\boldsymbol{X}$ is quasi-left continuous, and that $\boldsymbol{X}$ possesses the branching law if the Lévy system of the process $\boldsymbol{X}$ exists. Thus, in particular, we have the following

Theorem 1. 4. *If a branching Markov process $\boldsymbol{X}$ is a Hunt process with a reference measure,*[41] *then $\boldsymbol{X}$ is given as an (X^0, π)-branching Markov process.*

The construction of an (X^0, π)-branching Markov process for the given X^0 and π will be discussed in Chapter III and Chapter IV.

§1.3. M-equation and S-equation

Let $\boldsymbol{X}=(\boldsymbol{X}_t, \boldsymbol{P}_{\boldsymbol{x}})$ be a branching Markov process on $\hat{S}$. Set, for $f \in \boldsymbol{B}(S)$,

$$u(t, \boldsymbol{x}) = \boldsymbol{T}_t f(\boldsymbol{x}),$$

then $u(t, \boldsymbol{x})$ satisfies

$$u(t, \boldsymbol{x}) = \boldsymbol{T}_t^0 f(\boldsymbol{x}) + \int_0^t \int_{\boldsymbol{S}} \psi(\boldsymbol{x};\ dr\, d\boldsymbol{y}) u(t-r, \boldsymbol{y}), \tag{1.39}$$

where

$$\boldsymbol{T}_t^0 f(\boldsymbol{x}) = \boldsymbol{E}_{\boldsymbol{x}}[f(\boldsymbol{X}_t);\ t<\tau], \tag{1.40}$$

and

$$\psi(\boldsymbol{x};\ dr\, d\boldsymbol{y}) = \boldsymbol{P}_{\boldsymbol{x}}[\boldsymbol{X}_\tau \in d\boldsymbol{y}, \tau \in dr]. \tag{1.41}$$

Indeed, (1. 39) is a direct consequence of the strong Markov property (cf. Remark 0. 1 (ii)) applied to the Markov time τ.

Now we assume $\boldsymbol{X}$ is an (X^0, π)-branching Markov process, then by (1. 38)

$$\psi(x;\ dr\, d\boldsymbol{z}) = \int_S K(x;\ dr\, dy) \pi(y, d\boldsymbol{z}), \quad x \in S, \tag{1.42}$$

41) Cf. §1.1, footnote (14).

where $K(x: dr\, dy)$ is a kernel on $S\times([0,\infty)\times S)$ given by

(1.43) $$K(x;\, dr\, dy)=P_x^0[\zeta^0\in dr,\ X^0_{\zeta^0-}\in dy].$$

Therefore, if we set

(1.44) $$F(x;f)=\int\pi(x,d\boldsymbol{z})\widehat{f}(\boldsymbol{z}),\qquad f\in \boldsymbol{B}^*(S),$$

then (1.39) can be written in the form

(1.45) $$u(t,x)=T_t^0f(x)+\int_0^t\int_S K(x;\, dr\, dy)F(y;\, u(t-r,\cdot)),\qquad x\in S.$$

Here, we have used the branching property $u(t,\boldsymbol{x})=\widehat{u(t,\cdot)|_S}(\boldsymbol{x})$.

We shall call the equation (1.39) M-*equation.* It is an equation on $\boldsymbol{S}$ and it holds for any strong Markov process (but, of course, if $\boldsymbol{X}$ is a branching Markov process, $\boldsymbol{T}_t^0$ and Ψ possess certain structure, namely the Property B. III). The equation (1.45) will be called *S-equation.* It is a non-linear equation on S. The detailed study of these equations will be given in part III, Chapter IV.

References

[1] Bartlett, M. S.; An introduction to stochastic processes with special reference to methods and applications. Cambridge University Press, 1955.

[2] Bharucha-Reid, A. T.; Elements of the theory of Markov processes and their applications. McGraw-Hill, 1960.

[3] Bourbaki, N.; Topologie générale, Hermann, 1961.

[4] Courrège, Ph. et. P. Priouret; Temps d'arrêt d'une fonction aléatoire: Relations d'équivalence associées et propriétés de décomposition. Publi. Inst. Statist. Univ. Paris **14** (1965), 245–274.

[5] Courrège, Ph. et P. Priouret; Recollements de processus de Markov, Publi. Inst. Statist. Univ. Paris **14** (1965), 275–377.

[6] Dynkin, E. B.; Markov processes, Springer 1965.

[7] Feller, W.; On boundaries and lateral condititions for the Kolmogorov differential equations. Ann. Math. **65**. (1957) 527–570.

[8] Harris, T. E.; The theory of branching processes, Springer, 1963.

[9] Hille, E. and R. S. Phillips; Functional analysis and semigroups. Colloq. Publ. Amer. Math. Soc. 1957 2nd. ed.

[10] Hunt, G.; Markoff processes and potentials I, II. III. Ill. Jour. Math. **1** (1957) 44–93; **1** (1957), 316–369; **2** (1958), 151–213.

[11] Ikeda, N., M. Nagasawa and K. Sato; A time reversion of Markov processes with killing. Kōdai Math. Sem. Rep. **16** (1964), 88–97.

[12] Ikeda, N, M. Nagasawa and S. Watanabe; On branching Markov processes,

Proc. Japan Acad. **41** (1965), 816-821.
[13] Ikeda, N, M. Nagasawa and S. Watanabe; Fundamental equations of branching Markov processes, Proc. Japan Acad. **42** (1966), 252-257.
[14] Ikeda, N, M. Nagasawa and S. Watanabe; A construction of Markov processes by piecing out, Proc. Japan Acad. **42** (1966), 370-375.
[15] Ikeda, N, M. Nagasawa and S. Watanbe; A construction of branching Markov processes, Proc. Japan Acad. **42** (1966), 380-384.
[16] Ikeda, N, M. Nagasawa and S. Watanabe; Transformation of branching Markov processes, Proc. Japan Acad. **42** (1966), 719-724.
[17] Ikeda, N, M. Nagasawa and S. Watanabe; On branching semi-groups I, II. Proc. Japan Acad. **42** (1966), 1016-1021; **42** (1966), 1022-1026.
[18] Ikeda, N. and S. Watanabe; On some relation between the harmonic measures and Lévy measures for a certain class of Markov processes, Jour. Math. Kyoto Univ. **2**. (1962), 79-95.
[19] Itô, K. and H. P. McKean, Jr.; Diffusion processes and their sample paths. Springer 1965.
[20] Itô, K. and S. Watanabe; Transformation of Markov processes by multiplicative functionals, Ann. Inst. Fourier **15** (1965), 13-30.
[21] Kato, T.; Nonlinear evolution equations in Banach spaces, Proc. of Symp. in Appl. Math. of Amer. Math. Soc. **17** (1965). 50-67.
[22] Kelley, J. L.; General topology. Van Norstrand, 1955.
[23] Knight, F. B.; Brownian random chain, (pre-print).
[24] Kolmogorov, A. and N. A. Dmitriev; Branching stochastic processes. Doklady Akad. Nauk SSSR **56** (1947), 5-8.
[25] Kolmogorov, A., I. Petrovsky, et N. Piscounoff; Etude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique, Bulletin de l'Université d'état à Moscou, **1** (1937). 1-25.
[26] Kunita, H.; Applications of Martin boundaries to instantaneous return Markov processes over a denumerable space, Jour. Math. Soc. Japan, **14**, (1962), 66-100.
[27] Kunita, H. and S. Watanabe; On square integrable martingales, Nagoya Jour. Math. **30** (1967) 209-245.
[28] Kunita, H. and T. Watanabe; Note on transformations of Markov processes connected with multiplicative functionals, Mem. Fac. Sci. Kyushu Univ. A **17** (1963), 181-191; Supplement to the same paper, the same journal **18** (1964), 114-117.
[29] Loève, M.; Probability theory, D. Van Nostrand, 1960.
[30] Meyer, P. A.; Fonctionnelles multiplicatives et additives de Markov, Ann. Inst. Fourier **12** (1962), 125-230.
[31] Meyer, P. A.; Probability and potentials, Blaisdell, 1966.
[32] Motoo, M. and S. Watanabe; On a class of additive functionals of Markov processes, Jour. Math. Kyoto Univ. **4** (1965), 429-469.
[33] Moyal, J. E.; Discontinuous Markov processes, Acta Math. **98** (1957), 221-264.
[34] Moyal, J. E.; The general theory of stochastic population processes. Acta. Math. **108** (1962), 1-32.
[35] Moyal. J. E.; Multiplicative population processes. Jour. Appl. Prob. **1** (1964), 267-283.
[36] Mullikin, T. W.; Limiting distributions for critical multitype branching processes with discrete time. Trans. Amer. Math. Soc. **106** (1963), 469-494.
[37] Nagasawa, M. and K. Sato; Some theorems on time change and killing of

Markov processes. Kōdai. Math. Semi. Rep. **15** (1963), 195–219.
[38] Ryser, H. J.; Combinatorial Mathematics, John Wiley & Sons 1963.
[39] Sato, K. and T. Ueno; Multi-dimensional diffusion and Markov process on the boundary. Jour. Math. Kyoto Univ. **4** (1965), 529–605.
[40] Segal, I. E.; Non-linear semi-groups. Ann. of Math. **78**, (1963). 339–364.
[41] Sevastjanov, B. A.; Branching stochastic processes for particles diffusing in a restricted domain with absorbing boundaries. Theory of Prob. & its Appl. **3** (1958), 121–136.
[42] Silverstein, M. L.; Markov processes with creation of particles Z. Wahrsch. **9** (1968), 235–257.
[43] Skorohod, A. V.; Branching diffusion processes. Theory of Prob. & its appl. **9** (1964) 492–497.
[44] Volkonsky, V. A.; Additive functionals of Markov processes. Trudy Moskov Mat. Obšč. **9** (1960) 143–189.
[45] Watanabe, S.; On discontinuous additive functionals and Lévy measures of a Markov process. Japanese Jour. Math. **34**, (1964), 53–70.
[46] Watanabe S.; On the branching process for Brownian particles with an absorbing boundary. Jour. Math. Kyoto Univ. **4** (1965), 385–398.
[47] Watanabe, T.; A remark on the strong Markov property. Mem. Fac. Sci. Kyushu Univ. A. **17** (1963), 176–180.
[48] Yosida, K.; Functional Analysis. Springer, 1965.

OSAKA UNIVERSITY
TOKYO INSTITUTE OF TECHNOLOGY
AND KYOTO UNIVERSITY